Klaus Miesenberger
Joachim Klaus
Wolfgang Zagler
Arthur Karshmer (Eds.)

LNCS 4061

Computers Helping People with Special Needs

10th International Conference, ICCHP 2006
Linz, Austria, July 2006
Proceedings

Springer

Lecture Notes in Computer Science 4061

Commenced Publication in 1973
Founding and Former Series Editors:
Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Klaus Miesenberger   Joachim Klaus
Wolfgang Zagler   Arthur Karshmer (Eds.)

Computers Helping People with Special Needs

10th International Conference, ICCHP 2006
Linz, Austria, July 11-13, 2006
Proceedings

Springer

Volume Editors

Klaus Miesenberger
University of Linz
Institute "Integriert Studieren"
Altenberger Strasse 69, 4040 Linz, Austria
E-mail: klaus.miesenberger@jku.at

Joachim Klaus
University of Karlsruhe (TH)
Study Center for Blind and Partially Sighted Students (SZS)
Engesserstr. 4, 76128 Karlsruhe, Germany
E-mail: klaus@szs.uni-karlsruhe.de

Wolfgang Zagler
Vienna University of Technology
Institute "Integriert Studieren"
Favoritenstr. 9/029, 1040 Wien, Austria
E-mail: zagler@fortec.tuwien.ac.at

Arthur Karshmer
University of South Florida
Department of Information Technology
3433 Winter Lake Road, Lakeland, FL 33803, USA
E-mail: arthur@lklnd.usf.edu

Library of Congress Control Number: 2006928507

CR Subject Classification (1998): H.5.2, H.5.3, H.3, H.4, K.4, K.3

LNCS Sublibrary: SL 3 – Information Systems and Application, incl. Internet/Web and HCI

ISSN 0302-9743
ISBN-10 3-540-36020-4 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-36020-9 Springer Berlin Heidelberg New York

Springer is a part of Springer Science+Business Media

springer.com

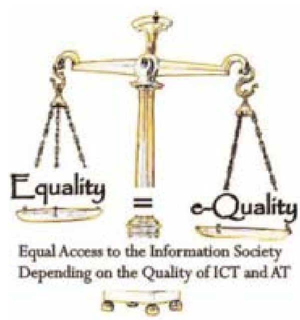
Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper SPIN: 11788713 06/3142 5 4 3 2 1 0

Preface

Welcome to the 10th ICCHP 2006!

Since the 1980s, computer technology has become a key technology for people with disabilities, offering functionalities to overcome barriers and take part in more and more areas of society. Starting with handicraft and trials, a new research and development field has been established known today as 'assistive technologies' (AT). Based on standard information and communication technology (ICT) the field has increasingly made impact on mainstream areas and establishes 'e-accessibility', 'e-inclusion' and 'design for all' as key concepts for an open and democratic society. The progress in ICT and AT significantly pushes forward the disability rights and independent living movement. The obvious benefits for the target group help to understand the way we design the emerging information society, which determines accessibility for and inclusion of people with disabilities.



The motto of the 10th ICCHP, 'equality = e-Quality: Equal Access to the Information Society Depending on the Quality of ICT and AT,' reflects this development. ICCHP's mission is to support this process. ICCHP has been one of the first meeting places for pioneers establishing the field. It supports today the interdisciplinary discussion and presentation of research and development which is inevitable when aiming at better quality of life for people with disabilities. ICCHP is proud of its role. With a focus on scientific excellence, ICCHP has become an important reference to AT for the mainstream. After more than two decades the first steps have been made, but lot more is needed to make the information society a really inclusive one. ICCHP is committed to supporting this development in the future. Come and get involved!

ICCHP 2006 was held under the auspices of Dr. Heinz Fischer, President of the Federal Republic of Austria. ICCHP 2006 was supported by several sponsors, to name only a few:

- European Commission, DG Information Society, eInclusion
- Federal Ministry for Education, Science and Culture
- Federal Ministry for Transport, Innovation and Technology
- Regional Government Upper Austria, Governor
- Regional Government Upper Austria, Department for Social Affairs
- Austrian Computer Society

Our special thanks go to those contributing toward this conference:

General ICCHP 2006 Chair, H. J. Murphy, Founder, Former Director and Member Advisory Board of the Centre on Disabilities, California State University, Northridge, USA

Programme Chairs

Karshmer, A.	University of South Florida at Tampa	USA
Burger, D.	INSERM	France
Suzuki, M.	Kyushu University	Japan
Tjoa, A. M.	Vienna University of Technology	Austria
Vollmar, R.	Universität Karlsruhe (TH)	Germany
Wagner, R.	University of Linz	Austria

Publishing Chairs

Klaus, J.	Universität Karlsruhe (TH)	Germany
Miesenberger, K.	University of Linz	Austria
Zagler, W.	Vienna University of Technology	Austria

Young Researchers Consortium Chairs

Alm, N.	Dundee University	UK
Archambault, D.	Université Piere et Marie Curie	France
Blenkhorn, P.	University of Manchester	UK
Evans, G.	University of Manchester	UK
Fitzpatrick, D.	Dublin City University	Ireland
Kobayashi, M.	Tsukuba College of Technology	Japan
Pontelli, E.	New Mexico State University	USA
Sinclair, R.	Microsoft	USA
Weber, G.	Multimedia Campus Kiel	Germany

Programme Committee

Abascal, J.	Euskal Herriko Univertsitatea	Spain
Abou-Zahra, S.	World Wide Web Consortium (W3C)	France
Andrich, R.	SIVA Fondazione Don Carlo Gnocchi Onlus	Italy
Arató, A.	KFKI-RMKI	Hungary
Arató, P.	TU Budapest	Hungary
Azevedo, L.	Instituto Superior Tecnico	Portugal
Batusic, M.	University of Linz	Austria
Bothe, H.-H.	Technical University of Denmark	Denmark
Brown, D.	The Nottingham Trent University	UK
Bühler, C.	FTB	Germany
Carves, J.	IBM	France
Charlier, J.	Metrovision	France
Craddock, G.	Central Remedical Clinic Dublin	Ireland

Crombie, D.	FNB	The Netherlands
Cummins Prager, M.	California State University Northridge	USA
Darzentas, J.	University of the Aegean	Greece
Dotter, F.	Universität Klagenfurt	Austria
Edwards, A.	University of York	UK
Ekberg, J	Stakes	Finland
Emiliani, P.L.	Istituto di Fisica Applicata 'Nello CARRARA'	Italy
Engelen, J.	Katholieke Universiteit Leuven	Belgium
Evreinov, G.	University of Tampere	Finland
Fels, D	Ryerson University	Canada
Freitas, D.	University of Porto	Portugal
Gardner, J.	Oregon State University	USA
Gupta, G.	University of Texas at Dallas	USA
Holzinger, A.	Graz University Hospital	Austria
Ioannidis, G	University Bremen	Germany
Knops, H.	iRv	The Netherlands
Kowalik, R.	Technical University of Gdansk	Poland
Kremser, W.	OCG	Austria
Lauruska, V.	Siauliai University	Lithuania
Leahy, D.	Trinity College Dublin	Ireland
Magnussen, M.	THF	Sweden
McKenzey, N.	FNB	The Netherlands
McMullin, B.	Dublin City University	Ireland
Mendelova, E.	Comenius University of Bratislava	Slovak Republic
Mokhtari, M.	Inst. Nat. Supérieur des Télécommunications	France
Neveryd, H.	Lund University	Sweden
Nicolle, C.	Loughborough University	UK
Nischelwitzer, A.	FH Johaneum	Austria
Nussbaum, G.	KI-I	Austria
Ohnuma, N.	Tsukuba College of Technology	Japan
Ohtake, N.	Tsukuba College of Technology	Japan
Ono, T.	Tsukuba University of Technology	Japan
Ortner, D.	University of Linz	Austria
Oussik, V.	Logos	Russia
Paciello, M.	The Paciello Group	USA
Panek, P.	Technical University of Vienna	Austria
Penaz, P.	University of Brno	Czech Republic
Petrie, H	University of York	UK
Petz, A.	University of Linz	Austria
Prazak, B.	Austrian Research Centres	Austria
Pühretmair, F.	KI-I	Austria
Quirchmayr, G.	University of Vienna	Austria

Raisamo, R.	University of Tampere	Finland
Schweikhardt, W.	University of Stuttgard	Germany
Schwerdtfeger, R	IBM	USA
Soede, T.	iRv	The Netherlands
Stephanidis, C.	ICS-Forth, Greece	Greece
Stoeger, B.	University of Linz	Austria
Strauss, Ch.	University of Vienna	Austria
Svensson, H.	National Institute for Special Needs Education	Sweden
Tahkokallio, P.	Stakes	Finland
Traunmüller, R.	University of Linz	Austria
Trehin, P.	World Autism Organisation	Belgium
Tretter, F.	KH Haar	Germany
Vaspöri, T.	KFKI-RMKI	Hungary
Velasco, C.	Fraunhofer Institute	Germany
Velleman, E.	Bartimeus	The Netherlands
Wagner, G.	University of Applied Sciences Steyr	Austria
Wöß, W.	University of Linz	Austria
Wünschmann, W.	TU Dresden	Germany

Organizing Committee (University of Linz)

Arrer, B.	Batusic, M.	Bogner, T.	Fahrengruber, C.
Feichtenschlager, P.	Hengstberger, B.	Herramhof, S.	Matausch, K.
Miesenberger, K.	Ortner, D.	Petz, A.	Recknagl, N.
Ruemer, R.	Schult, C.	Schüler, C.	Stöger, B.

We thank the Austrian Computer Society for announcing and sponsoring the **Roland Wagner Award on Computers Helping People with Special Needs**.

The Austrian Computer Society decided in September 2001 to endow this award in honor of Prof. Dr. Roland Wagner, the founder of ICCHP.

The Roland Wagner Award is a biannual award and will be in the range of 3000 euro. It will be handed over at the occasion of ICCHP conferences.

Award Winners:

- Award 0 was handed over to Prof. Dr. Roland Wagner on the occasion of his 50th birthday.
- Award 1 was won by WAI-W3C. It was handed over to Judy Brewer at ICCHP 2002.
- Special Award 2003: A special award was handed over to Prof. Dr. A. Min Tjoa, one of the founders of ICCHP, on the occasion of his 50th birthday in 2003.
- Award 2 was won by Paul Blenkhorn, University of Manchester, and handed over during ICCHP 2004 in Paris.

Once again we thank all those helping to put ICCHP together and thereby supporting the AT field and a better quality of life for people with disabilities.

July 2006

Klaus Miesenberger
Joachim Klaus
Arthur Karshmer
Wolfgang Zagler
(Editors)

Table of Contents

People with Disabilities: Accessible Content Processing

People with Disabilities: Accessible Content Processing
David Crombie 1

HODDER – A Fully Automatic Braille Note Production System
Matthias Leopold 6

Maps Sonification System Using Digitiser for Visually Impaired Children
Gintautas Daunys, Vidas Lauruska 12

Document Processing for Accessibility: Standards and Initiatives
George Ioannidis, David Crombie, Francisco Martínez Calvo, Konstantina N. Geramani 16

Expressing Emotions Using Animated Text Captions
Raisa Rashid, Jonathan Aitken, Deborah I. Fels 24

Schulbuch Barrierefrei (Accessible School Books) – Co-operation Between Publishers and Service Providers in Austria
Klaus Miesenberger, Reinhard Ruemer 32

Modelling Accessibility Constraints
Terje Gjøsæter, Jan Pettersen Nytun, Andreas Prinz, Mikael Snaprud, Merete Skjelten Tveit 40

Scientific PDF Document Reader with Simple Interface for Visually Impaired People
Toshihiro Kanahori, Masakazu Suzuki 48

English "Oblique" Listening System – Rapid Listening System for the Blind and Visually Impaired, and Its Evaluation
Shinichi Torihara, Miyoko Nakamura, Nami Ueda, Takuya Wada, Shun Ishizaki 53

Creating Pictures by Dialogue
Ivan Kopeček, Radek Ošlejšek 61

Automatic Annotation of Geographic Maps
Mirko Horstmann, Wilko Heuten, Andrea Miene, Susanne Boll 69

Towards the Use of Ontologies for Improving User Interaction for People with Special Needs
Shuaib Karim, A. Min Tjoa 77

A Model for Accessible Information Networks – Findings of the EUAIN Project
Benoît Guillon, Catherine Desbuquois, Dominique Burger 85

Making the Past a Thing of the Future: Automated Workflow for the Conversion of Printed Items into Fully Structured Digital Objects Based on Common Open Metadata Standards
Claus Gravenhorst 92

Accessible Navigation of Rich Media: Exposing Structure, Content and Controls in the Mobile User Interface
Markku T. Häkkinen 96

The Essential Role of Libraries Serving Persons Who Are Blind and Print Disabled in the Information Age
George Kerscher 100

People with Disabilities: Web Accessibility

Browsing Web Based Documents Through an Alternative Tree Interface: The WebTree Browser
Esmond Walshe, Barry McMullin 106

Web Pages for Blind People — Generating Web-Based Presentations by Means of Dialogue
Luděk Bártek, Ivan Kopeček 114

Architecture for Personal Web Accessibility
Myriam Arrue, Markel Vigo, Julio Abascal 120

Dynamically Generated Scalable Vector Graphics (SVG) for Barrier-Free Web-Applications
Kerstin Altmanninger, Wolfram Wöß 128

Designing a Web Page Considering the Interaction Characteristics of the Hard-of-Hearing
Miki Namatame, Tomoyuki Nishioka, Muneo Kitajima 136

Accessible Websites for People with Dementia: A Preliminary Investigation into Information Architecture
Nada Savitch, Panayiotis Zaphiris 144

People with Disabilities: Automatic and Manual Evaluation of Websites

People with Disabilities: Automatic and Manual Evaluation of Web Sites
Helen Petrie, Gerhard Weber 152

Web Accessibility Testing: When the Method Is the Culprit
Giorgio Brajnik 156

Test Case Description Language (TCDL): Test Case Metadata for Conformance Evaluation
Christophe Strobbe, Sandor Herramhof, Evangelos Vlachogiannis, Carlos A. Velasco 164

The BenToWeb XHTML 1.0 Test Suite for the Web Content Accessibility Guidelines 2.0
Christophe Strobbe, Sandor Herramhof, Evangelos Vlachogiannis, Johannes Koch, Carlos A. Velasco 172

An Environment for Defining and Handling Guidelines for the Web
Barbara Leporini, Fabio Paternò, Antonio Scorcia 176

Interpreting Results from Large Scale Automatic Evaluation of Web Accessibility
Christian Bühler, Helmut Heck, Olaf Perlick, Annika Nietzio, Nils Ulltveit-Moe 184

Managing and Monitoring Web Site Accessibility
Shadi Abou-Zahra 192

Semi-automatic Evaluation of Web Accessibility with HERA 2.0
Carlos Benavídez, José L. Fuertes, Emmanuelle Gutiérrez, Loïc Martínez 199

Automatic Control of Simple Language in Web Pages
Constantin Jenge, Sven Hartrumpf, Hermann Helbig, Rainer Osswald 207

Test Case Management Tools for Accessibility Testing
Sandor Herramhof, Helen Petrie, Christophe Strobbe, Evangelos Vlachogiannis, Kurt Weimann, Gerhard Weber, Carlos A. Velasco 215

Testing Web Navigation for All: An Agent-Based Approach
Martin González, Marcos González, Cristóbal Rivera, Ignacio Pintado, Agueda Vidau 223

University of Illinois Tools and Techniques for Functional Web Accessibility
Hadi Bargi Rangin 230

A Proposed Architecture for Large Scale Web Accessibility Assessment
Mikael Holmesland Snaprud, Nils Ulltveit-Moe, Anand Balachandran Pillai, Morten Goodwin Olsen 234

People with Disabilities: Quality of Web Accessibility

People with Disabilities: Quality of Web Accessibility
Dominique Burger 242

Towards Web Accessibility Certification: The Findings of the Support-EAM Project
Dominique Burger, Pierre Guillou 245

Raising the Expertise of Web Designers Through Training – The Experience of BFWD – Accessible Web Design (Barrierefreies Webdesign) in Austria
Klaus Miesenberger, Daniela Ortner 253

The Role of Benchmarking in Concerted Actions to Increase Accessibility
Finn Aslaksen, Frank Fardal, Mikael Snaprud 258

A Ubiquitous Social Community Portal Service for Social Networking with Convenient Accessibility
Yung Bok Kim 263

Web Accessibility Conformity Assessment – Implementation Alternatives for a Quality Mark in Austria
Marie-Luise Leitner, Klaus Miesenberger, Daniela Ortner, Christine Strauss 271

Accessible Interfaces to Empower the User
Mari Luz Guenaga, Javier Oliver, Ander Barbier 279

A GOMS Model for Keyboard Navigation in Web Pages and Web Applications
Martin Schrepp, Patrick Fischer 287

People with Disabilities: Accessible Tourism

People with Disabilities: Accessible Tourism
Franz Pühretmair 295

Accessibility Add-on Box Enabling Barrier-Free Tourism Information Systems (TIS)
Michael Winkler, Wolfram Wöß 298

Development of a Mobile Tourist Information System for People with Functional Limitations: User Behaviour Concept and Specification of Content Requirements
Sascha M. Sommer, Marion Wiethoff, Sari Valjakka, Dionisis Kehagias, Dimitrios Tzovaras 306

Objective Accessibility Assessment of Public Infrastructures
Christophe Ponsard, Vincent Snoeck 314

BAIM–Information for People with Reduced Mobility in the Field of Public Transport
Christian Bühler, Helmut Heck, Dirk Sischka, Josef Becker 322

Accessible Information Space to Promote Accessible Tourism
Franz Pühretmair, Klaus Miesenberger 329

People with Disabilities: Materials for Teaching Accessibility and Design for All

People with Disabilities: Materials for Teaching Accessibility and Design for All
Gerhard Weber, Julio Abascal 337

Teaching Web Accessibility with “Contramano” and Hera
Carlos Benavídez, José L. Fuertes, Emmanuelle Gutiérrez, Loïc Martínez 341

Introducing Media Managers to Usability and Accessibility
Christoph Haffner, Gerhard Weber 349

Evaluating the Length of Virtual Horizontal Bar Chart Columns Augmented with Wrench and Sound Feedback
Tatiana G. Evreinova, Grigori Evreinov, Roope Raisamo 353

“Assistec” – A University Course on Assistive Technologies
Kerstin Matausch, Barbara Hengstberger, Klaus Miesenberger 361

People with Disabilities: Entertainment Software Accessibility

People with Disabilities: Entertainment Software Accessibility
Dominique Archambault 369

Semi Automatic Generator of Tactile Video Games for Visually Impaired Children
Alexis Sepchat, Nicolas Monmarché, Mohamed Slimane, Dominique Archambault 372

Making the Mainstream Accessible: What's in a Game?
Matthew T. Atkinson, Sabahattin Gucukoglu, Colin H.C. Machin, Adrian E. Lawrence 380

Access Invaders: Developing a Universally Accessible Action Game
Dimitris Grammenos, Anthony Savidis, Yannis Georgalis, Constantine Stephanidis 388

Internet and Accessible Entertainment
Morten Tollefsen, Are Flyen 396

Guidelines for the Development of Accessible Computer Games
Roland Ossmann, Klaus Miesenberger 403

People with Disabilities: Human Computer Interface

A Classification, Based on ICF, for Modelling Human Computer Interaction
M. Billi, L. Burzagli, P.L. Emiliani, F. Gabbanini, P. Graziani 407

Evaluation of Reaction Forces During Human Computer Interaction for Optimization and Development - A Pilot Research
Stefan Mina, Lisa Ehrenstrasser, Robert Lurf, Barbara Prazak 415

A User-Orientation Evaluation Framework: Assessing Accessibility Throughout the User Experience Lifecycle
Alexandros Mourouzis, Margherita Antona, Evagelos Boutsakis, Constantine Stephanidis 421

Basic Research of Input Support Device by Using Sympathetic Skin Response
Chikamune Wada, Fumihiko Masuda 429

Unskilled Finger Key Pressing and Brain Coherence
Ling-Fu Meng, Chiu-Ping Lu, Ching-Horng Chen 437

Customizing User Interfaces with Input Profiles
Mario Mühlehner 442

Computerized Assessment Approach for Evaluating Computer Interaction Performance
Ming-Chung Chen, Chi-Nung Chu, Ting-Fang Wu, Chih-Ching Yeh 450

People with Disabilities: Assistive Homes and Environments

People with Disabilities: Assistive Homes and Environments
Gerhard Nussbaum 457

Prototyping and Evaluation of New Remote Controls for People with Visual Impairment
Michiaki Yasumura, Ryo Yoshida, Megumi Yoshida 461

Constructing Adaptive User Interface Through Semantic Descriptions
Hyun Namgoong, Kyung-il Kim, Yun Koo Chung 469

A Proposal for a Home-Based Health Monitoring System for the Elderly or Disabled
HyungJun Kim, Bart Jarochowski, DaeHyun Ryu 473

Helping People with ICT Device Control by Eye Gaze
Fangmin Shi, Alastair Gale, Kevin Purdy 480

IRCS – Infra Red Code System: Access to Infra Red Codes for Ambient and Assisted Living
Klaus Miesenberger, Gerhard Nussbaum, Martin Bartsch, Thomas Mayrhofer, Christoph Strobl-Mairhofer 488

The DAT Project: A Smart Home Environment for People with Disabilities
Renzo Andrich, Valerio Gower, Antonio Caracciolo, Giovanni Del Zanna, Marco Di Rienzo 492

Selection by Visualization of Topological Layouts for Adapted Living Area Design
Arnaud Puret, Sébastien Aupetit, Pierre Gaucher, Nicolas Monmarché, Mohamed Slimane 500

Wireless Communicator for Patients in Intensive Care
Cecília Sik Lányi, Viktor Magyar 508

People with Disabilities: Service Delivery

The Development and Application of the Assistive Devices Application System for Clients with Disabilities
Tzyh-Chyang Chang, Jiann-Der Lee, Shwu-Jiuan Wu, Juei-Fen Huang 516

EASTIN: A Trans-national Information Network on Assistive Technologies
Renzo Andrich, Roberto Da Dalt, Renzo Giust, Andrea Agnoletto . . . 524

Participation in Development of Computers Helping People
Lars Bruhn, Jürgen Homann, Gerlinde Renzelberg 532

Developing Pedagogical Multimedia Resources Targeting Children with Special Educational Needs
Paloma Cantón, Ángel Lucas González, Gonzalo Mariscal, Carlos Ruiz 536

People with Disabilities: Education and Training

Accessibility Issues in Two Specifications for E-Learning Tests: IMS QTI 1.2 and IMS QTI 2.0
Christophe Strobbe 544

Informatics Teachers and Their Competences in Inclusive Education
L'udmila Jašková 552

ECDL bf: Equal Opportunities Through Equal Access to an ECDL E-Learning Solution
Andrea Petz, Klaus Miesenberger 560

Searching Knowledge CinemaSense as a Case Study in Collaborative Production of a WWW Service in Two Universities
Antti Raike 568

Deaf and Hard of Hearing People: Electronic Communication Aids

Deaf and Hard of Hearing People: Electronic Communication Aids
Hans-Heinrich Bothe 575

User Evaluation of the SYNFACE Talking Head Telephone
Eva Agelfors, Jonas Beskow, Inger Karlsson, Jo Kewley, Giampiero Salvi, Neil Thomas 579

Towards a Service Integration Portal for Deaf People
Christophe Ponsard, Christiane Broekman, Cécile Lamy, Martine Fraiture 587

A Proposal of the Universal Sign Code
Kazuyuki Kanda, Tsutomu Kimura, Daisuke Hara 595

Bilingual Sign Language Dictionary
José L. Fuertes, Ángel L. González, Gonzalo Mariscal, Carlos Ruiz 599

Improvements and Evaluations in Sign Animation Used as Instructions for Stomach X-Ray Examination
Kazunari Morimoto, Takao Kurokawa, Syouhei Kawamura 607

Speech Technologies in a Computer-Aided Speech Therapy System
András Kocsor, Dénes Paczolay 615

Automatic Synthesis of Training Data for Sign Language Recognition Using HMM
Kana Kawahigashi, Yoshiaki Shirai, Jun Miura, Nobutaka Shimada 623

Communication Supporting System in a Classroom Environment for the Hearing Impaired
Yoshinori Takeuchi, Yudai Sakashita, Daisuke Wakatsuki, Hiroki Minagawa, Noboru Ohnishi 627

Efficient Generation of Large Amounts of Training Data for Sign Language Recognition: A Semi-automatic Tool
Ruiduo Yang, Sudeep Sarkar, Barbara Loeding, Arthur Karshmer 635

Composition Corrector – A Computer-Based Tool to Support Professional Deaf and Hard-of-Hearing Writers
Martha Birnbaum, Corine Bickley, Brianne Burger 643

VoxAid 2006: Telephone Communication for Hearing and/or Vocally Impaired People
Bálint Tóth, Géza Németh 651

Evaluation of Effect of Delay on Sign Video Communication
Kaoru Nakazono, Yuji Nagashima, Mina Terauchi 659

Japanese JSL Translation and Searching Display Conditions for Expressing Easy-to-Understand Sign Animation
Sumihiro Kawano, Chie Izumi, Takao Kurokawa, Kazunari Morimoto 667

Design and Development of Several Mobile Communication Systems for People with Hearing Disabilities
Jose L. Martín, Sira E. Palazuelos, Jerónimo Arenas, Javier Macías, Santiago Aguilera 675

Captioning for Deaf and Hard of Hearing People by Editing Automatic Speech Recognition in Real Time
Mike Wald 683

Improving Baby Caring with Automatic Infant Cry Recognition
Sandra E. Barajas-Montiel, Carlos A. Reyes-García, Emilio Arch-Tirado, Mario Mandujano 691

People with Cognitive Problems and the Aging Population

People with Cognitive Problems and the Aging Population
Erwin Fugger, Martin Morandell, Barbara Prazak 699

A Semi-autonomous Wheelchair Towards User-Centered Design
Sven Rönnbäck, Jouni Piekkari, Kalevi Hyyppä, Tomas Berglund, Simo Koskinen 701

Computer Control by Tracking Head Movements for the Disabled
HyungJun Kim, DaeHyun Ryu 709

Healthcare Service with Ubiquitous Sensor Networks for the Disabled and Elderly People
Yung Bok Kim, Daeyoung Kim 716

Guiding Support for 'Way-Finding' in Unknown Buildings: Design and Evaluation
José M. Falcó, Roberto Casas, Álvaro Marco, Jorge L. Falcó 724

Mobile Computing in Medicine: Designing Mobile Questionnaires for Elderly and Partially Sighted People
Andreas Holzinger, Peter Sammer, Rainer Hofmann-Wellenhof 732

NeurOSS — Open Source Software for Neuropsychological Rehabilitation
Jyrki Rissanen 740

DALMA – Location Aware Alarm System for People with Disabilities
Roberto Casas, Álvaro Marco, Jorge L. Falcó, Héctor Gracia, José I. Artigas 744

Usage of IT and Electronic Devices, and Its Structure, for Community-Dwelling Elderly
Madoka Ogawa, Hiroki Inagaki, Yasuyuki Gondo 752

A Randomized, Placebo-Controlled Study of the Efficacy of Cognitive Intervention on Elderly People and on Patient's with Alzheimer's Disease
J.J. Yanguas, C. Buiza, I. Etxeberria, N. Galdona, M.F. González, E. Urdaneta 759

Compensatory Use of Computers by Disabled Older Adults
Boaz Kahana, Eva Kahana, Loren Lovegreen, Gul Seçkin 766

Computer Aids Clients with Psychiatric Disabilities in Cognitive and Vocational Rehabilitation Programs
Tzyh-Chyang Chang, Jiann-Der Lee, Shwu-Jiuan Wu, Ming-Jen Yang, Chun-Hua Shih, Juei-Fen Huang 770

Ethically Aware Design of a Location System for People with Dementia
Roberto Casas, Álvaro Marco, Jorge L. Falcó, José I. Artigas, Julio Abascal 777

People with Specific Learning Difficulties

TouchStory: Towards an Interactive Learning Environment for Helping Children with Autism to Understand Narrative
Megan Davis, Kerstin Dautenhahn, Chrystopher Nehaniv, Stuart D. Powell 785

Computer-Assisted Language Learning (CALL) for Dyslexic Students
Cara Nicole Greene 793

A New Audio Testing System for the Newly Blind and the Learning Disabled to Take the National Center Test for University Admissions
Mamoru Fujiyoshi, Akio Fujiyoshi 801

Embodied Agents in Language Learning for Children with Language Challenges
Dominic W. Massaro 809

Web Design for Dyslexics: Accessibility of Arabic Content
Areej Al-Wabil, Panayiotis Zaphiris, Stephanie Wilson 817

Flexibility in Virtual Environments – A Fully Adjustable Virtual Classroom
Zoltán Geiszt, Cecília Sik-Lányi, Péter Károlyi 823

Temporal Orientation Panel for Special Education
Jorge L. Falcó, Carmen Muro, Inmaculada Plaza, Armando Roy 831

Developing a TriAccess Reading Environment
Chien-Chuan Cko, Ming-Chung Chen, Li-Yuan Chen, Jun-Nan Guo, Ji-Fu Liu, An-Jun Luo 839

A Platform for Creating Adaptive Communicators
María Dolores Paredes Garrido, Oscar Pino Morillas, María José Rodríguez Fórtiz, Manuel González González, Encarnación Rodríguez Parra 847

Do Text-to-Speech Synthesisers Pronounce Correctly? A Preliminary Study
D.G. Evans, E.A. Draffan, A. James, P. Blenkhorn 855

People Using Augmented and Alternative Communication (AAC)

Investigation on Effect of Prosody in Finger Braille
Manabi Miyagi, Masafumi Nishida, Yasuo Horiuchi, Akira Ichikawa 863

Internet Chat System for the Deaf-Blind Using Doubled Braille Display – DB4DB
Makoto Kobayashi 870

Using Iconicity to Evaluate Symbol Use
David Gareth Evans, Lisa Bowick, Marianne Johnson, Paul Blenkhorn 874

Development of a Voice-Input Voice-Output Communication Aid (VIVOCA) for People with Severe Dysarthria
Mark S. Hawley, Pam Enderby, Phil Green, Stuart Cunningham, Rebecca Palmer 882

EMU – A European Multilingual Text Prediction Software
Christian Beck, Gottfried Seisenbacher, Georg Edelmayer, Wolfgang Zagler 886

Design and Evaluation of a Versatile Architecture for a Multilingual Word Prediction System
Sira E. Palazuelos-Cagigas, José L. Martín-Sánchez, Lisset Hierrezuelo Sabatela, Javier Macías Guarasa 894

People with Motor and Mobility Impairement: Human Computer Interaction, Rehabilitation

People with Motor and Mobility Impairment: Human Computer Interaction, Rehabilitation
Michiaki Yasumura, Kazuyuki Itoh, Akira Ichikawa, Akira Okamoto 902

Electromyogram-Based Cursor Control System for Users with Motor Disabilities
Craig Chin, Armando Barreto, Miguel Alonso Jr. 905

A Tongue Based Control for Disabled People
Lotte N.S. Andreasen Struijk 913

An Alternative Chinese Keyboard Layout Design for Single-Digit Typists
Ming-Chung Chen, Ting-Fang Wu 919

An Integrated Design for a Myoelectrically-Based Writing Module for a Controlled Prosthesis
Andres Herrera, Malek Adjouadi, Melvin Ayala 926

Development of a Power Assisted Handrail – Handrail Trajectory and Standing Up Motion
Yoshiyuki Takahashi, Osamu Nitta, Shigeru Okikawa, Takashi Komeda 935

Trial Usage of Rehabilitation System: Simple Driving Simulator for the Driving Skill Evaluation of People with Cerebrovascular Disease: CVD
Kaoru Inoue, Yuko Ito, Yumi Ikeda, Atsuko Tanimura, Keisuke Suzuki, Yoshiyuki Takahashi, Tadashi Suzuki, Madoka Fujimoto, Masahiro Tsuchijima, Takashi Komeda 943

Eye Movement Communication Control System Based on EOG and Voluntary Eye Blink
Junichi Hori, Koji Sakano, Michio Miyakawa, Yoshiaki Saitoh 950

Eye as an Actuator
Marcela Fejtová, Jan Fejt, Olga Štěpánková 954

Nonlinear Mapping of Pupil Centre Coordinates from Image Sensor to Screen for Gaze Control Systems
Gintautas Daunys, Nerijus Ramanauskas 962

Keyboard Adaptations for Children with Cerebral Palsy
Ting-Fang Wu, Ming-Chung Chen 966

Light Spot Operated Mouse Emulator for Cervical Spinal-Cord Injured PC Users
Kazuyuki Itoh 973

Design and Implementation of a Chorded On-Screen Keyboard for People with Physical Impairments
Yun-Lung Lin, Ming-Chung Chen, Yao-Ming Yeh, Wen-Jeng Tzeng, Chih-Ching Yeh 981

People with Motor and Mobility Impairement: Innovative Interfaces to Wheelchairs

People with Motor and Mobility Impairment: Innovative Multimodal Interfaces to Wheelchairs
Andreas Holzinger, Alexander K. Nischelwitzer 989

Accessible User Interface Framework for Severely Physically Disabled People
Michael Mahr, Alexander K. Nischelwitzer 992

MediaWheelie – A Best Practice Example for Research in Multimodal User Interfaces (MUIs)
Alexander K. Nischelwitzer, Bernd Sproger, Michael Mahr, Andreas Holzinger 999

NIBLUM: Non-Invasive Bluetooth Mouse for Wheelchair Users
Roberto Casas, Marcos Quilez, Borja Romero, Oscar Casas 1006

Smart Wheelchair Based on Ultrasonic Positioning System
HyungJun Kim, DaeHyun Ryu 1014

Scenarios of Use for a Modular Robotic Mobility Enhancement System for Profoundly Disabled Children in an Educational and Institutional Care Environment
Peter Mayer, Paul Panek, Georg Edelmayer, Marnix Nuttin, Wolfgang Zagler 1021

Blind and Visually Impaired People: Human Computer Interface

Blind and Visually Impaired People: Human Computer Interface
Grigori Evreinov 1029

A Framework for Blind User Interfacing
Fernando Alonso, José L. Fuertes, Ángel L. González, Loïc Martínez 1031

An Approach for Direct Manipulation by Tactile Modality for Blind Computer Users: Development of the Second Trial Production
Shigenobu Shimada, Masami Shinohara, Yutaka Shimizu, Makoto Shimojo 1039

Making Nonaccessible Applications Accessible for Visually Impaired
Tomáš Zahradnický, Róbert Lórencz 1047

Usability Evaluation of the MOST Mobile Assistant (SlatTalker)
Zoltan Juhasz, Andras Arato, Gabor Bognar, Laszlo Buday, Gergely Eberhardt, Norbert Markus, Emil Mogor, Zoltan Nagy, Terez Vaspori 1055

Non-visual Access to GUIs: Leveraging Abstract User Interfaces
Kris Van Hees, Jan Engelen 1063

Design Guidelines for Audio–Haptic Immersive Applications for People with Visual Disabilities
Fernando Alonso, José L. Fuertes, Loïc Martínez, Héctor Szabo 1071

The Amodal Communication System Through an Extended Directional Input
Georgios Yfantidis, Grigori Evreinov 1079

Blind and Visually Impaired People: Access to Information and Communication

Multimedia Browser for Internet Online Daisy Books
Piotr Brzoza, Dominik Spinczyk 1087

Automated Book Reader for Persons with Blindness
Malek Adjouadi, Eddy Ruiz, Lu Wang 1094

RoboBraille – Automated Braille Translation by Means of an E-Mail Robot
Lars Ballieu Christensen 1102

Accessing Music Notation Through Touch and Speech
Ben P. Challis 1110

Blind People: Access to Graphics

Math Class: An Application for Dynamic Tactile Graphics
Peter Albert 1118

Helping People with Visual Impairments Gain Access to Graphical Information Through Natural Language: The ***iGraph*** System
Leo Ferres, Avi Parush, Shelley Roberts, Gitte Lindgaard 1122

Fingertip Guiding Manipulator: Haptic Graphic Display for Mental Image Creation
Yoshihiko Nomura, Yuki Yagi, Tokuhiro Sugiura, Hirokazu Matsui, Norihiko Kato 1131

Graphic Editor for Visually Impaired Users
Atsushi Nishi, Ryoji Fukuda 1139

Exploiting SenseCam for Helping the Blind in Business Negotiations
Shuaib Karim, Amin Andjomshoaa, A. Min Tjoa 1147

On the Accuracy of Tactile Displays
Christopher Power 1155

HOWARD: High-Order Wavefront Aberration Regularized Deconvolution for Enhancing Graphic Displays for Visually Impaired Computer Users
Miguel Alonso Jr., Armando Barreto, Malek Adjouadi, Julie A. Jacko 1163

Blind People: Access to Mathematics

"BlindMath" a New Scientific Editor for Blind Students
A. Pepino, C. Freda, F. Ferraro, S. Pagliara, F. Zanfardino 1171

Translating MathML into Nemeth Braille Code
Paul B. Stanley, Arthur I. Karshmer 1175

New Environment for Visually Disabled Students to Access Scientific Information by Combining Speech Interface and Tactile Graphics
Toshihiko Komada, Katsuhito Yamaguchi, Fukashi Kawane, Masakazu Suzuki 1183

Canonical MathML to Simplify Conversion of MathML to Braille Mathematical Notations
Dominique Archambault, Victor Moço 1191

Mathematics: How and What to Speak
D. Fitzpatrick 1199

SBT: A Translator from Spanish Mathematical Braille to MathML
Fernando Alonso, José L. Fuertes, Ángel L. González, Loïc A. Martínez 1207

Braille Math Made Easy with the Tiger Formatter
John A. Gardner, Leon Ungier, John J. Boyer 1215

LAMBDA: A European System to Access Mathematics with Braille and Audio Synthesis
Waltraud Schweikhardt, Cristian Bernareggi, Nadine Jessel, Benoit Encelle, Margarethe Gut 1223

Speech Recognition Helps Visually Impaired People Writing Mathematical Formulas
Tomáš Hanakovič, Marek Nagy 1231

Supporting Blind Students in Navigation and Manipulation of Mathematical Expressions: Basic Requirements and Strategies
Bernhard Stoeger, Mario Batusic, Klaus Miesenberger, Philipp Haindl 1235

Scientific Diagrams Made Easy with IVEO^{TM}
John A. Gardner, Vladimir Bulatov 1243

Blind and Visually Impaired People: Mobility and Orientation

CyARM: Interactive Device for Environment Recognition and Joint Haptic Attention Using Non-visual Modality
Tetsuo Ono, Takanori Komatsu, Jun-ichi Akita, Kiyohide Ito, Makoto Okamoto 1251

Marking in the Surroundings by Data-Carriers for the Visually Impaired
Hisayuki Tatsumi, Yasuyuki Murai, Nobuyuki Nagai, Masahiro Miyakawa 1259

A Study and Development of the Auditory Route Map Providing System for the Visually Impaired
Takafumi Ienaga, Michito Matsumoto, Masahiro Shibata, Nobuyuki Toyoda, Youko Kimura, Hiroshi Gotoh, Takehito Yasukouchi 1265

Identification of Acoustic Factors for Perception of Crossability Common to Blind and Sighted Pedestrians
Takayuki Shiose, Kiyohide Ito, Kazuhiko Mamada 1273

A Computer-Based Navigation System Tailored to the Needs of Blind People
Bettina Pressl, Manfred Wieser 1280

A Haptic Interface for an Indoor-Walk-Guide Simulator
Yasuyuki Murai, Hisayuki Tatsumi, Nobuyuki Nagai, Masahiro Miyakawa 1287

Computer Vision-Based Terrain Sensors for Blind Wheelchair Users
James Coughlan, Roberto Manduchi, Huiying Shen 1294

The Blind Interactive Guide System Using RFID-Based Indoor Positioning System
Jongwhoa Na 1298

Blind and Visually Impaired People: Education and Training

An IT Training Programme for Blind Computer Users – Presentation and Discussion of Didactic and Teletutorial Implications
Mario Batušić, Andrea Gaal, Joachim Klaus, Mary O'Grady 1306

Usability for All: Towards Improving the E-Learning Experience for Visually Impaired Users
Eva Patrícia Gil Rodríguez, Muriel Garreta Domingo, Jordi Planella Ribera, Magí Almirall Hill, Llorenç Sabaté Jardí 1313

Issues in Implementing Awareness in Collaborative Software for Blind People
Jaime Sánchez, Nelson Baloian 1318

Modeling 3D Interactive Environments for Learners with Visual Disabilities
Jaime Sánchez, Nelson Baloian 1326

APL: Audio Programming Language for Blind Learners
Jaime Sánchez, Fernando Aguayo 1334

eDiab: A System for Monitoring, Assisting and Educating People with Diabetes
L. Fernández-Luque, J.L. Sevillano, F.J. Hurtado-Núñez, F.J. Moriana-García, F. Díaz del Río, D. Cascado 1342

Author Index 1351

People with Disabilities: Accessible Content Processing
Introduction to the Special Thematic Session

David Crombie

FNB Amsterdam,
Molenpad 2, The Netherlands
dcrombie@fnb.nl
http://www.euain.org

Abstract. The Special Thematic Session (STS) on Accessible Content Processing is intended to provide a focus for several different activities which can be grouped under this area. The European Accessible Information Network (EUAIN) was established in order to bring together the different stakeholders in the accessible content processing chain and to build on common concerns. EUAIN has now completed a systemic overview of this area and will use this information to provide guidelines, training materials and input into standardisation activities. The papers in this STS address many of these issues.

1 Introduction

Given the widespread adoption of ICT within the publishing industries, there is a general interest in the creation and provision of well-formatted digital documents. For those people who are dependent on accessible information, this interest is of central importance, and it is this convergence of interests that offers exciting opportunities for these different stakeholders. This workshop will examine the ways in which this convergence is helping to build consensus and create new standards and technologies for the provision of information in formats, which are more accessible for everyone. This thematic session focuses on this emerging area and try to provide a context for many of the changes now being seen at social, organisational and technical levels throughout the publishing world. For those involved in producing information in alternative formats this raises some formidable challenges. As technologies converge, and accessibility is gradually incorporated into mainstream creation, production and distribution environments, what then is the role of specialist organisations? And what about the needs of end users?

2 The EUAIN Network

The EUAIN project (European Accessible Information Network) (www.euain.org) is funded by the eInclusion thread of the European Commission 6th framework IST programme and co-ordinated by FNB Amsterdam. EUAIN has four key objectives, namely:

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 1 – 5, 2006.

- To bring together all the players in the information provision and e-publishing chain in order to achieve the critical mass significantly to enhance the provision of accessible information at a European level
- To create a self-sustaining network which can offer the necessary expertise, shared knowledge, technological tools and distribution platforms to all those involved in the provision of accessible information and in particular to provide a collaborative platform for content creators to find information, tools, advice and solutions.
- To raise awareness and stimulate the adoption at local, regional, national and European levels of the emerging formats and standards for the provision of accessible information and to find ways of ensuring that technological protection measures do not inadvertently impede legitimate access to information by people with print impairments.
- To stimulate and support the adoption of new distribution channels and appropriate business solutions in order to help achieve faster, prompter and less expensive production and distribution of accessible information

3 Special Thematic Session

The special thematic session on accessible content processing attempts to cover several key areas, namely:

- Modelling accessible content processing
- Structuring information
- Creating workflows for content creation/production/consumption
- Technical protection measures and copyright issues
- Enabling technologies, standards and tools
- Integrating accessible processing within mainstream environments
- Benefits of approach: accessibility on demand

This diverse range of topics represent the emerging themes that have come to the attention of the EUAIN network as the key areas of discussion around the subject of Accessible Content Processing. The STS features several presentations that deal with tools and technologies for the creation and processing of accessible content. These technologies cover a diverse range of formats and technologies which when included in frameworks and environments for accessible content processing, provide a means of integrating these technologies with mainstream workflows.

4 Areas Covered by STS

Dominique Burger (BRAILLENET) proposes a model for Accessible Information Networks based on the experience of the “Serveur Hélène” project and on-going research within the EUAIN network. It outlines the main components of the model adopted which supports contractual policies with publishers, as well as technical tools for improving the accessibility of the files provided.

Deborah Fels (Ryerson University) presents a framework for enhancing captions that uses animation and a set of standard properties to express five basic emotions. Using an action research method, the framework was developed from a designer's interpretation and rendering of animated text captions for two content examples.

Mirko Horstmann (University of Bremen) describes an approach to generate semantic descriptions of entities in city maps so that they can be presented through accessible interfaces. The solution we present processes bitmap images containing city map excerpts. Regions of interest in these images are automatically extracted based on colour information and their geometric properties are subsequently determined. The result of this process is a structured description of these regions based on the Geography Markup Language (GML), an XML based format for the description of GIS data. This description can later serve as an input to innovative presentations of spatial structures using haptic and auditory interfaces.

Vidas Lauruska (Siauliai University) showcases a system that investigates graphical information for blind people using a digitiser. The SVG language with additional elements is used for describing of maps. Non-speech sounds are used to transfer information about colour. A sound alert is issued for the boundary between two regions.

Matthias Leopold (DZB) presents software that is able to generate Braille Music Notes fully automatically, thereby focusing on flexibility, readability, quality and translation speed. The software is designed to be used within a team of professional Braille note transcribers and to take over much of the tedious part of their work. Moreover the software can be used as a stand-alone product.

Shinichi Torihara (Keio University) proposes an "oblique" listening system in the English language for the blind and visually impaired by controlling the speed of Text-to-Speech based on parts of speech in which the important parts are synthesized 'relatively' slower while the unimportant at the maximum speed. In the system evaluation experiment, English-speaking natives with visual impairments were required to listen to three short passages of Text-to-Speech at the three types of speed and answer the questions of comprehension after the base-line speed was calculated from the measured maximum speed of recognizing 'a word' and that of recognizing 'a sentence' for each subject.

The STS also intends to place these technologies within a wider context. In order to do this, the STS also provides examples of endeavours where such integration has already taken place.

Claus Gravenhorst (CCS Content Conversion Specialists GmbH) discusses the basic need for accessible information. The METAe Project and its results are presented, as related to the digitisation of printed materials, the advantages of harvesting structural metadata during the process, an overview of the digitisation technology and workflow in the docWORKS conversion solution from CCS.

Klaus Miesenberger(University of Linz) outlines the co-operation between Austrian school book publishers and service providers for people with special needs which aims at making books available in electronic formats. In this project the following has been established: 1. a minimum set of structural elements which documents from publishers have to contain to make them usable for the production of books in alternative formats (e.g. Braille, large print, ebooks) 2. know-how and handouts for publishers on how to implement structured design with these elements using standard desktop publishing (DTP) systems (InDesign, QarkExpress) 3. examples of new books and re-designing existing books to learn how to perform accessible document design in practice 4. training materials, workshops and seminars to transfer the developed know-how to other publishers and design agencies 5. a general agreement which gives the right of transferring books in electronic format to students with disabilities 6. a Document Rights Management System including to prevent the data to be misused in practice 7. a workflow for the co-operation between schools/teachers, service providers, publishers and the ministry.

An important testbed for many of the technologies and practices described in the EUAIN network is that of educational and academic publishing.

Toshihiro Kanahori presents an overview of an integrated system for scientific documents including mathematical formulae with speech output interface, named ``ChattyInfty." This system consists of recognizer, editor and speech interface for scientific documents. Using this system, visually impaired people can read printed scientific documents with speech output. The author proposes a new additional function of this system which recognizes PDF documents. Using this function, visually impaired people can also read PDF documents including mathematical formulae with ChattyInfty. If PDF documents have embedded text, the PDF recognition engine does not only recognize PDF documents as page images but also utilizes the text information to verify its recognition results, to improve the recognition results.

One important area of the EUAIN network is the conformance with emerging and existing standards. George Ioannidis (TZI) provides an overview of the work being done by the Document Processing for Accessibility Workshop that is co-ordinated by CEN/ISSS and summarizes its findings so far. This focuses on standards and initiatives that have been elaborated and concerns document management and accessibility issues especially for the publishing industry. The paper also provides an outline of the planned work and invites interested parties to take an active role in shaping the workshop.

George Kerscher (DAISY consortium) provides an update on the "traditional" activities of libraries for the blind, and reveals the essential roles that many libraries for the blind are playing in the standards, open source software, and the consumer products arena that will lead to that vision of the future.

Some more generic presentations which directly address accessible information processing bring the various topics together in order better to illustrate how the subject can encompass the various requirements of all stakeholders represented and ensure that these requirements can be integrated and communicated.

Radek Oslejsek (Masaryk University) deals with the problem of generating a picture of a scene by means of describing it in terms of natural language ontologies. The corresponding graphical format, being formally represented in the form of the Pawlak information system, involves a full semantic description of the picture. Therefore, the pictures generated in this way are ``blind-friendly", i.e. they can be automatically fully described.

Mikael Snaprud (Agder University College) describes the combination of research topics from two projects focusing on web accessibility testing, and on metamodelling and high-level specification languages. This combination leads to a novel approach to accessibility assessment that should improve the understanding of accessibility issues, and explore the potential for generating executable accessibility tests based on accessibility constraint models.

The STS also hosted a keynote panel session to discuss the concerns emerging from the thematic session. A number of keynote speakers, including George Kersher (DAISY Consortium), Robert Martinengo (Centre for Accessible Publishing) and Rob Havertly (Microsoft) will participate. This panel discussion was entitled: *"Is there still a role for specialist accessible information providers?"*

More information about the work of EUAIN and all current materials are available from the EUAIN portal (`www.euain.org`).

Acknowledgements

The EUAIN project (European Accessible Information Network) (www.euain.org) is funded by the eInclusion thread of the European Commission 6th framework IST programme. The author would like to acknowledge the collective work of the project partners in making this session posssible.

HODDER – A Fully Automatic Braille Note Production System

Matthias Leopold

Deutsche Zentralbücherei für Blinde zu Leipzig (DZB)
Gustav-Adolf-Straße 7
04105 Leipzig
Germany
Tel.: +49 341 7113 190; Fax: +49 341 7113 125
info@dzb.de
http://www.dzb.de

Abstract. We present the software HODDER which is able to generate Braille Music Notes fully automatically. Thereby we focus on flexibility, readability, quality and translation speed. The software is designed to be used within a team of professional Braille note transcribers and to take over much of the tedious part of their work. Moreover the software is able to be used stand-alone.

1 Introduction

Braille music notation is a means of displaying music notes to blind persons. It has been invented by Louis Braille and is nowadays established as an international standard for music notation for the blind.

Because producing Braille scores is very costly, there has been nearly no Braille note production in Germany and some other countries for many years. However, in 2000, the German Central Library for the Blind in Leipzig has restarted to produce Braille notes.

In 2003, we started with DaCapo[1], a project which aims at speeding up the Braille score production by developing and using suitable software. Thus, we will establish an enduring Braille score production in Leipzig.

1.1 The Need for Automatic Note Transcription

The simplest way of producing Braille scores is to write them by hand. This is not the slowest way, but it is error-prone and the quality of its results depends on the qualification of the transcriber. Another problem is the inflexibility of the produced notes: imagine a piece of music which has to be transposed for some reason (for instance, someone might like to play Mozart's clarinet concerto – which is written for a clarinet in A – with his/her clarinet in B). Manual transposition involves a complete rewriting of the notes, whereas any sophisticated

[1] The project is externally financed by the German Federal Ministry for Labour and Social Affairs (http://www.bmas.bund.de).

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 6–11, 2006.

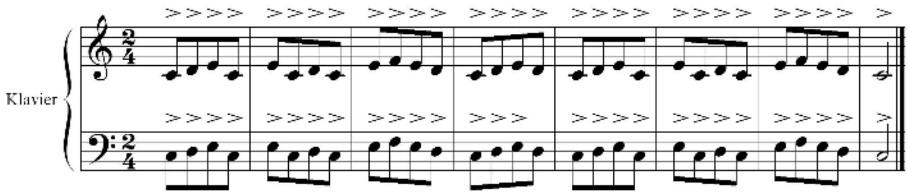



Fig. 1. Example on references, parallel shifts and short written accents: Translating literally from ink print results in 176 Braille forms. Inserting references and short written forms (specialities of Braille music notation) can reduce this to 29 Braille forms. Measures 5, 6, 7 are an exact copy of measures 1, 2, 3. This can be easily and very shortly expressed using Braille music notation. It is also possible to efficiently encode that *all* notes are played accentuated instead of attaching accents to each note. Only four braille forms can indicate to the blind musician that the left hand is playing the same as the right hand. This way he can save reading time and is supported in learning the notes by heart.

music software can transpose the notes in a matter of seconds and absolutely error-free – even after they have been transcribed to Braille.

1.2 The Challenge of Braille Note Production

Blind musicians have other needs than sighted ones, because they are not reading scores visually but using their fingers. Because of the limited information that can be perceived at a given moment, Braille music notation offers a variety of special notation forms (mostly abbreviations) to increase the readability of the notes. These notations are very different from standard music notation. See figure 1 for some examples.

A system which is designed to automatically translate ink printed notes into Braille notes will have to deal with this issue, i.e. it is absolutely crucial not to simply return a "literal translation" but to assure readability of the Braille notes by using the special notations.

The problem with this is, that – depending on the context – there are often different alternatives of choosing the exact notation form. The final choice depends on many factors such as the context of the music and a human transcriber's experience. It is thus not well defined and leaves room for interpretation. A translation software will have to incorporate much of a human transcriber's experience and flexibility in handling different contexts. In the rest of this paper, I will introduce Hodder, a system for automatic Braille note transcription with a special focus on readability and flexibility.

2 State of the Art

During the last few years, there have been some propositions on how to make music accessible to blind musicians.

Some of these were beyond the standard Braille music notation: Talking Music [9], the hapto-graphical representation of music [2] extended by the use of force feedback mouse [5] or the haptic printing of ink notes using wax, as proposed by the universities of Würzburg and Augsburg [4].

In the world of Braille, two main ideas of automatic Braille score production have emerged:

1. The first idea is to produce a one-to-one copy of the ink printed notes in Braille (i.e. a "literal translation") and then to let a professional note transcriber insert all Braille note specifics *afterwards*. This is not expected to be much faster than writing notes by hand nor is it much more flexible. It is, however, less error-prone since part of the work is done automatically. Installing the Braille note specifics can be supported by external software like the Braille Music Editor [3].
2. The second idea is to extend and modify that concept. The Braille specifics are not added *after* having set the notes but *before*. Therefore an ink print-like meta format is defined which allows to enter those specifications [7]. The focus of this approach is on controlability. There is no implementation yet, but the concept is more flexible than the first solution in that all translation-specific options will not be lost when the translation has to be repeated.

In the next sections, I will propose an extended solution combining elements of the ideas described above. The main intention is to maximize speed, quality, flexibility and – above all – readability of the produced notes.

3 Development of a Braille Music Meta Format

I have extended the data structure BrailleMusicXML proposed by Christian Waldvogel [6] by the possibilities, to represent both, ink print and braille music notes at the same time, while the original proposition was limited to represent ink print, only, providing possibilities to describe Braille like a super layer over the ink print. The reasons for our extensions are as follows

- We want to be able to transform Braille scores to ink print for services, music note verification and visualisation. Therefore we have to be able to represent pure Braille as well as ink print.
- We assume to be more independent of future developments if the main transformation processes can be done completly on BrailleMusicXML
- With respect to time limitations and capacities our proposition can be easier implemented than the second idea mentioned above.

4 The HODDER System

I have developed the software Hodder implementing above described concepts. Hodder is designed to be an autonome Braille score transcriber having strong points but also limits like all Braille score transcribers.

The goal is HODDER to be integrated into a team of professional Braille score transcribers. The (dis)abilities) of HODDER will be learned and understood by it's "Collegues". For that reason the confidence in HODDER will raise time by time and the production speed will increase because of saving time for verifying HODDER's output.

4.1 General Architecture

The core of HODDER's architecture is BrailleMusicXML as defined by Christian Waldvogel and extended by myself. Note transformation is understood as a logical process transforming an ink print close music representation to a Braille music close representation and vice versa. See figure 2. Apart of the core there are interfaces to different formats. Those formats can be added via interfaces to the HODDER system. Additional ink print formats could be MusicTEX, Midi or new developments. Additional Braille interfaces could produce different Braille music setting formats as well as library specifics.

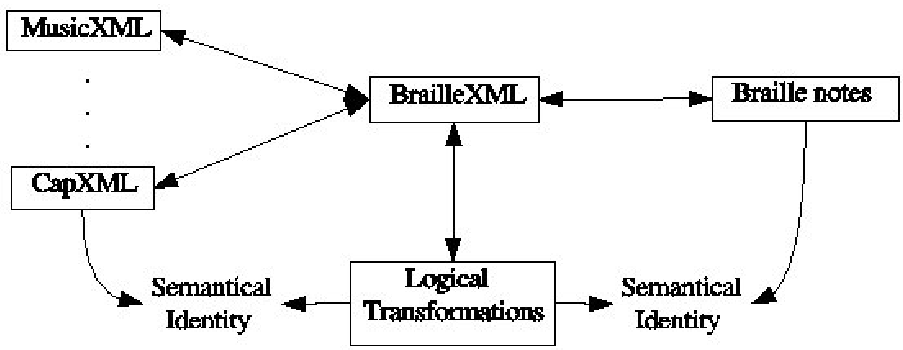


Fig. 2. Overview over the architecture of HODDER: The ink print formats CapXML and MusicXML as well as Braille are connected by interfaces to BrailleMusicXML while almost all transformation are happening on BrailleMusicXML itself

4.2 Transformation Rules

As explained above, the core of the system consists of logical transformation rules which ensure the readability of the Braille output by introducing special Braille notations and abbreviations (again, see figure (1) for some examples).

The rules for short Braille notations are not well defined within the Braille music notation manual [8], but rather embedded within examples reflecting special musical situations.

In addition, these examples are somewhat redundant because they contain many musical situations, which can, in principle, be handled in a very similar fashion, i.e. by the same set of rules. As an example, consider the following abbreviation candidates:

1. repetition of beats (simile) and other parts of a measure (simile)
2. repetition of single (simile) or more measures or voices (simile)
3. references within a musico-logical section or overall score
4. parallel shifts (left plays same as right hand)
5. sequence abbreviations (repeat but starting on another step)

All these situations can be handled by the following basic rule:

(Part1, CompressedPart2) = compress(Part1, Part2) ←
similar(Part1, Part2) ∧
CompressedPart2 = ... *% Compress Info Tag...*

(Part1, Part2) = compress(Part1, Part2).

This rule simply states that, if two parts (whatever their exact nature may be) are similar, the second part should be compressed. Of course the actual work will now be to construct a good and general similarity measure for the two parts.

For that reason, the first step in designing a transformation rule system must be to (conceptually) make the rules from the manual *explicit* so that they can be expressed by logical rules as indicated above.

Since there are often many ways to do so, it would be desirable to define standards on how Braille music scores should be written. These standards can and should be extended by options offering alternatives and exceptions. We have included alternatives in the Hodder system, in order to provide flexibility with respect to the transcribers' and musicians' whishes.

5 Findings

We are already using Hodder to set Braille notes. The software is continually tested and improved within a note production context. Because of the high quality output, the production speed is nearly doubled in comparison to manual transcription.

In the process of developing and using the system, a lot of interesting questions and problems have arisen. In the following, I will list a small selection of our findings:

1. *Controlability*: Interviews with users have revealed that they want to have as much control over options of the transformation process as possible. However, these options were not actually used much (or misused) in reality. For that reason, I have installed countless auto decision makers which take over control whenever the user does not. Thus, users have the feeling that they can control everything if they want to, but they are not obliged to do it.
2. *Limitations of Scanning Software*: The recognition of graphical objects often poses problems. The output of scanning software tools like Capella [1] often contain simple graphical objects instead of musico-logical notations (XML). Hodder is not always able to interpret those objects in a correct way automatically. We are, however, working together with the developers of the

OCR Software and have already made improvements to the specific needs of the translation software HODDER.
3. *Quality of output*: All in all, it can be said that the output of *inner content* (musical and musico-logical notations like notes, simile etc.) are nearly 95-100% error-free. *Outer content* (questions of arrangement, dispatching special words, different assignments) still causes problems. Therefore we are focussing more and more to those questions.

6 Future Work

1. In the future we want to be able to accept not only scanned files but also the sources of the publishers. Therefore we will have to extend the range of ink print input files (primarily MusicXML). Furthermore, the software has to learn to cope with incorrect input (auto corrections, hints to the user etc.)
2. Out of whole partitures we want to be able to automatically extract voices for the different musicians. Imagine a choir where the different singers soprano, alto, tenore and basso need their own notes while the choir leader needs all notes in a special, different format, the score format.
3. In opposition to ink print the readability (speed and understanding) of Braille scores is dependent of the involved informations. E.g. while fingering in ink print could easily be ignored they could use an amount of space and reading time in Braille scores and also complicating learning by heart in cases when they are not needed. We want to offer notes to the specials needs of the blind musician which we call notes on demand.

References

1. Capella. `http://www.capella.de/`.
2. J.M. Chouvel, G. Assayag, and J.P. Cholleton. La reprsentation hexagonale torodale, application l'analyse harmonique de la musique d'hermeto pascal. In *Sminaire MaMux*, 2002.
3. Braille Music Editor. `http://www.dodiesis.com/`.
4. Frank Giener and Wolfgang Weber et al. Forschungsprojekt "PRINT". `http://test.fh-augsburg.de/projekte/blindenschriftdrucker/`.
5. Nadine Baptiste Jessel. Access to musical information. In *International Symposium on Braille Music Notes (Leipzig)*, 2005.
6. Thomas Kahlisch, Matthias Leopold, and Christian Waldvogel. Dacapo, a project on transforming ink printed to braille notes semi-automatically. In *Proc. of Computers Helping People with Special Need; 9th International Conference*, 2004.
7. Matthias Leopold. Development of a DTD for representing braille music. In *International Symposium on Braille Music Notes (Leipzig)*, 2005.
8. Braille Music Notation Manual. `http://www.brl.org/music/`.
9. Talking Music. `http://projects.fnb.nl/Talking%20Music/`.

Maps Sonification System Using Digitiser for Visually Impaired Children

Gintautas Daunys and Vidas Lauruska

Siauliai University, 76353, 141 Vilniaus Str., Siauliai, Lithuania
{g.daunys, vidas.lauruska}@tf.su.lt

Abstract. Presentation of graphical information is key problem during teaching of visually impaired children. The developed system is devoted for investigation of graphical information by blind user using a digitiser. SVG language with additional elements is used for describing of maps. Non-speech sounds are used to transfer information about colour. Alerting sound signals is issued near two regions boundary.

1 Introduction

Visualisation is commonly used within almost every scientific field. For teaching blind children or students other information presentation ways must be found, which replace visual information. Solution is to transform visual information to stimulus which could be perceived by others human sensor systems, which are functioning normally. We selected approach that transformation to auditive signals must be preferred. The aim of this study is to create cheap system for maps sonification.

1.1 Visual Information Transformation

There are many attempts to develop systems for blind people to explore graphical objects. Tactile (embossed) maps were designed for this purpose. Until recently, the use of tactile maps has been very limited, as most of the maps have been produced by hand, for example using string and glue. Recent developments facilitated the production of cost effective maps. For examples: printers, new types of papers and new types of ink.

Other approach is based on transformation of visual information to auditive signal. As example is coding scheme based on a pixel-frequency association [1]. In agreement with cochlear tonotopy, sinewave frequencies associated to pixels should be equally distributed along an exponential scale to be perceived as approximately equidistant.

In the TeDUB project ("Technical Drawings Understanding for the Blind")[2] the system was developed, which aim is providing blind computer users with an accessible representation of technical diagrams. The TeDUB system consists of two separate parts: one for the semi-automatic analysis of images containing diagrams from a number of formally defined domains and one for the representation of previously analysed material to blind people. The joystick was used for navigation through drawings. Very interesting approach and ideas are combining haptic and auditory [3].

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 12 – 15, 2006.

1.2 SVG Format

Nowadays vector graphics format is widely used to store digitized maps. Often rich interactive maps are published in web using SVG [4] file format. SVG (abbreviation for Scalable Vector Graphics) is an XML markup language for describing two-dimensional vector graphics. It is an open standard created by the World Wide Web Consortium. The available fill and stroke options, symbols and markers enable higher quality map graphics.

Most suitable software for browsing interactive SVG maps is plugin Adobe SVG Viewer, available for all major platforms and browsers (Linux, MacOSX, Solaris, Windows) which can be downloaded free from the Adobe SVG homepage [5]. Exist and commercial products as MapViewSVG from ESRI [6].

The analysis of SVG technology for application in education is presented A. Neumann paper [7]. Already were attempts to apply SVG formats of maps for blind users[8].

2 Method

Our aim was to develop widely available graphical information presentation system for blind user. We tried to use most common and cheapest hardware and open source or free software components.

First we consider the system hardware. Computer mouse is optional graphic-input device. The device use relative motion, so when the user hits the edge he or she need merely pick up the mouse and drop it back. It is convenient during usual work with computer applications, but maps exploration system is one of exceptions. In our application we need devices which give absolute coordinates. For graphical input we selected digitiser (tablet). Other hardware is a standard PC computer with soundcard and Microsoft Windows operation system and load speakers.

For graphical information description we selected SVG language because of reasons as was described earlier. Because we aimed achieved high interactivity, we don't use standard products with SVG implementation as Adobe SVG Reader. We developed software using Visual C++ environment from Microsoft Visual Studio.NET.

As a XML based language, SVG supports foreign namespaces. It is possible to define new elements or add new attributes. Elements and attributes in a foreign namespace have a prefix and a colon before the element or attribute name. Elements and attributes in foreign namespaces that the SVG viewer does not know, are ignored. However, they can be read and written by script. Foreign namespaces are used to introduce new elements (e.g. GUI elements, scalebars) and for the attachment of non-graphical attributes to SVG graphic elements [7]. The SVG file, prepared for our system could be seen on other SVG viewers.

3 Maps Sonification System

Important and coherent points during design process were to choose structure for parsed SVG code storing and define additional elements to SVG specification. The hierarchical structure for maps information storage is shown in Figure 1. Entire document is represented as node Map. Some text information could be assigned to entire map. Map consists of regions. Every region is described by external contour with its attributes, as line colour and filling colour. For each region musical melodies are prescribed. The sound is produced when digitizer pen is over selected region. Optional text field could contain information for about the region. The information is presented for user by synthesised voice, when user gives command to listen attached to region information.

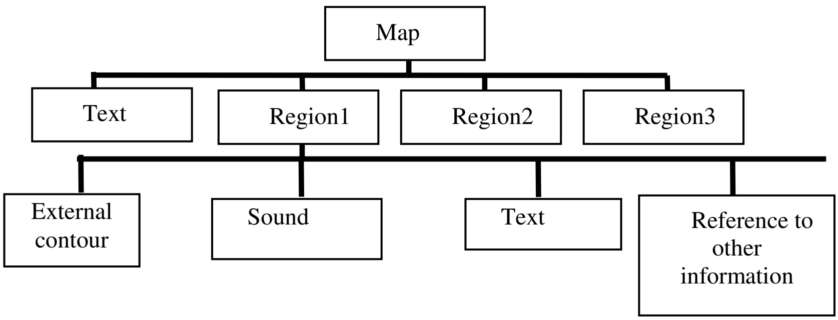


Fig. 1. Hierarchical structure for parsed information storage

For easier navigation through the map, the alerting system about region boundaries was implemented. System permanently tracks, how far user input coordinates are from region contours. In region near boundaries an alert signal is issued. When user input is closer to boundary then the volume of signal is increased.

The parsing of SVG document was implemented using XML parser Expat [9]. Graphical rendering was implemented with Windows GDI functions. We used DirectX for non-speech sound output. Speech synthesis currently is implemented using free product Microsoft SAPI 5. For us seems that application became more attractive, when we tonical sounds replaced by short musical melodies. The WAV file could be assigned during editing of SVG file. In other case the melodies are selected from application library.

4 Conclusions

In the paper we have demonstrated that SVG format became important for implementation of interactive maps. The software prototype, which uses SVG format files as input was implemented using Microsoft Visual Studio.NET. Usability test is our nearest future task. To improve speech synthesis quality and include language choice menu is other task. It seems that MBROLA tool [10] is suitable for the purpose.

Acknowledgement. This research is supported by FP6/IST STREP project Multimodal Collaboration Environment for Inclusion of Visually Impaired Children (MICOLE).

References

1. Capelle, C., Trullemans, C., Amo, P., Veraart, C.: A Real-Time Experimental Prototype for Enhancement of Vision Rehabilitation Using Auditory Substitution. IEEE Trans. Biom. Eng. BME-45 (1998) 1279-1293
2. Horstmann, M., Hagen, C., King, A., Dijkstra, S., Crombie, D., Evans, D.G., Ioannidis, G., Blenkhorn, P., Herzog, O., Schlieder, Ch.: TeDUB: Automatic Interpretation and Presentation of Technical Diagrams for Blind People. In Marion Hersh (eds.), CVHI 2004 (2004)
3. Jansson, G., Larsson, K.: Identification of haptic virtual objects with different degrees of complexity. In S. A. Wall, B. Riedel, A Crossan & M. R. McGee (Eds.), Eurohaptics 2002, Conference Proceedings, Edinburgh, (2002) 57-60
4. Scalable Vector Graphics (SVG) 1.1 specification. `http://www.w3.org`
5. Adobe downloads page. `http://www.adobe.com/support/downloads/main.html`
6. ESRI homepage. `http://www.esri.com`
7. Neumann, A.: Use of SVG and EcmaScript technology for E-learning purposes. In: ISPRS Workshop Tools and Techniques for E-Learning, Potsdam, Germany, June 1-3 2005
8. Campin, B., McCurdy, W., Brunet, L., Siekierska, E.: SVG Maps for People with Visual Impairment. In SVG OPEN Conference, July 2003
9. The Expat XML parser. http://expat.sourceforge.net
10. The MBROLA project homepage. `http://tcts.fpms.ac.be/synthesis`

Document Processing for Accessibility: Standards and Initiatives

George Ioannidis[1], David Crombie[2], Francisco Martínez Calvo[3], and Konstantina N. Geramani[4]

[1] Technologie-Zentrum Informatik (TZI), University Bremen, Am Fallturm 1, D-28359 Bremen, Germany
george.ioannidis@tzi.de

[2] FNB Amsterdam
Molenpad 2, 1016GM, Amsterdam, The Netherlands
projects@fnb.nl

[3] ONCE Madrid
Prim, 3, 28004, Madrid, Spain
fmc@once.es

[4] TWT GmbH, Information & Engineering Technologies
Bernhäuser Str. 40-42, D-73765 Neuhausen, Germany
konstantina.geramani@twt-gmbh.de

Abstract. This paper presents the Document Processing for Accessibility Workshop that is held at CEN/ISSS and summarizes its findings so far. It focuses on standards and initiatives that have been elaborated and concern document management and accessibility issues especially for the publishing industry. The paper provides also an outline of the work to be done and invites interested parties to take an active role in shaping the workshop.

1 Introduction

The Workshop (WS) on Document Processing for Accessibility (DPA) is an initiative carried out in the framework of the European Committee for Standardization (CEN) and its domain of Information Society Standardization System (ISSS) [5]. The DPA workshop was initiated by the EUAIN [1] (European Accessible Information Network) project. EUAIN is a Co-ordination Action project (contract number 511497) co-funded by the INFSO DG of the European Commission within the RTD activities of the Thematic Priority Information Society Technologies of the 6th Framework Programme.

Given the widespread adoption of Information and Communication Technologies (ICT) within the publishing industries, there is a general interest in the creation and provision of well-formatted digital documents. For those people who are dependent on accessible information, this interest is of central importance, and it is this convergence of interests that offers exciting opportunities for these different stakeholders.

CEN/ISSS WS/DPA will examine the ways in which this convergence is helping to build consensus and create new standards and technologies for the provision of

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 16 – 23, 2006.

information in formats that are more accessible for everyone. CEN/ISSS WS/DPA has three key objectives [2], namely:

1. To bring together all the players in the information provision and e-publishing chain in order to achieve the critical mass significantly to enhance the provision of accessible information at a European level.
2. To provide guidelines on integrating accessibility components within the document management and publishing process rather than as just a specialized, additional service.
3. To raise awareness and stimulate the adoption at local, regional, national and European levels of the emerging formats and standards for the provision of accessible information and to find ways of ensuring that technological protection measures do not inadvertently impede legitimate access to information by people with print impairments.

The DPA workshop has at the moment around 50 members, and additional 40 persons interested in its outcomes. The readers interested in taking an active role in the DPA workshop are invited to register at: http://www.cenorm.be/cenorm/business-domains/businessdomains/isss/activity/ws-dpa.asp. One of its first tasks was the elaboration of standards and initiatives for document management and accessibility [3] that are relevant for the publishing industry and the publishing process. This elaboration was based on a questionnaire that was distributed to the members and the interested parties and will be presented in the next sections. Before that, we will present why standards are important in the field and at the end draw our conclusions and highlight issues that have emerged.

2 Why Standards Are Important and Useful?

Standards are needed for many reasons, but probably the most relevant one is that they tell manufacturers how to make their products accessible in a detailed, coherent way. Legislations promote the existence of standards, and they advocate for "accessible" technology or information. But it's standards that give the technical specifications of how this accessibility can be implemented and tested.

The existence of standards though does not imply that accessibility will be implemented in the same way or with the same results in all products. The existence of a number of standards for producing the same product (a document) may occasionally lead to two different levels of accessibility for the same "accessible" final product.

It is also important to note that the higher the level of accessibility we want to apply to a document, the higher the cost. Therefore, it is of great importance to decide beforehand the level of accessibility to apply to a certain document according to different variables: (a) the depth of the structure that the document allows/needs to make it sufficiently navigable; (b) the level of navigability actually needed by the potential user/s; and (c) the resources available to make a document accessible.

One can distinguish between formal (or de jure) standards and de facto standards. The former are those who have been "formalized" by standards organizations, while the latter are technical solutions that have been adopted informally by users due to its usefulness and/or reliability. Among these de facto standards, we also have two categories – proprietary standards and open standards (outside vendors' control, freely developed and updated by independent programmers).

CEN/ISSS WS/DPA relies on the existence and the promotion of accessibility standards to prove that accessibility can be built from the first stage of production, that Design For All can be applied to emerging standards so that all the features needed to grant accessibility to the final product are built into the system right from the beginning, instead of the traditional approach of adding those features later.

3 Relevant Standards and Initiatives

In this section the relevant existing standards and activities gathered so far with the help of a questionnaire to interested parties and members of the DPA workshop and can be used to make content accessible will be briefly presented. For a more comprehensive analysis the reader is referred to [3].

3.1 ANSI/NISO Z39.86 (DAISY) and NIMAS

This standard for creating digital content in structured multimedia is developed and maintained by the DAISY Consortium [10]. Using XML text files and MP3 audio files, with the DAISY format we can create a range of text only, fully synchronized text and audio and audio only books that are fully accessible and navigable for blind and visually impaired users as well as persons with other disabilities, such as dyslexia. The DAISY version that is now being used more widely is 2.02, though NISO has already ratified the 2005 revision of Z39.86, which will become version 3.00.

The DAISY format was first aimed to substitute analogue "talking books". Today, the standard allows for different types of books, so that they can include not only the audio rendition of the printed text but the text itself and/or images to go with the text. A sub-set of the DAISY standard was proposed as the National File Format and in 2004 it was declared a "National Instructional Materials Accessibility Standard" (NIMAS).

3.2 WAI (Web Accessibility Initiative)

The Web is an increasingly important resource in many aspects of life: education, employment, government, commerce, health care, recreation, and more. It is essential that the Web be accessible in order to provide equal access and equal opportunity to people with disabilities. An accessible Web can also help people with disabilities more actively participate in society.

The Web offers the possibility of unprecedented access to information and interaction for many people with disabilities. That is, the accessibility barriers to print, audio, and visual media can be much more easily overcome through Web technologies.

One of the roles of the Web Accessibility Initiative (WAI) [21] is to develop guidelines and techniques that describe accessibility solutions for Web software and Web developers. These WAI guidelines are considered the international standard for Web accessibility.

3.3 Web Content Accessibility Guidelines (WCAG 1.0)

These guidelines [17] became a W3C recommendation on May 5th, 1999. There's already a working draft for WCAG 2.0 [18]. They explain how to make Web content accessible to people with disabilities and are intended for all Web content developers (page authors and site designers) and for developers of authoring tools. The primary goal of these guidelines is to promote accessibility. However, following them will also make Web content more available to all users, whatever user agent they are using (e.g., desktop browser, voice browser, mobile phone, automobile-based personal computer, etc.) or constraints they may be operating under (e.g., noisy surroundings, under- or over-illuminated rooms, in a hands-free environment, etc.).

3.4 WCAG 2.0 (Web Content Accessibility Guidelines 2.0)

Web Content Accessibility Guidelines 2.0 (WCAG 2.0) [23] covers a wide range of issues and recommendations for making Web content more accessible. This document contains principles, guidelines, success criteria, benefits, and examples that define and explain the requirements for making Web-based information and applications accessible. "Accessible" means to a wide range of people with disabilities, including blindness and low vision, deafness and hearing loss, learning difficulties, cognitive limitations, limited movement, speech difficulties, and others. Following these guidelines will also make your Web content more accessible to the vast majority of users, including older users. It will also enable people to access Web content using many different devices - including a wide variety of assistive technology.

3.5 AIIM PDF-Universal Access Working Group

AIIM (The Association for Information and Image Management) holds a number of standards committees and working groups that draft recommended practices for different activities. These drafts are being reviewed and revised until agreement to be submitted to ANSI for approval. The PDF-Universal Access Working Group [4] deals specifically with how PDF documents can be fully accessible. The starting point is that PDF files can be accessible, they just need to incorporate a number of guidelines on how to convey the information that traditionally has been only useful for sighted users.

3.6 CEN/ISSS Dublin Core Metadata Workshop (MMI-DC)

The Dublin Core is already a de jure standard in the United States of America, under the name ANSI/NISO Z39.85. The CEN/ISSS Dublin Core Metadata Workshop [6] will pave the way for this standard to be also considered as such in Europe, where is already being used as a de facto standard.

3.7 CEN/ISSS Design for All Workshop and Assistive Technologies in ICT

There are two main means of ensuring that people with disabilities benefit from ICT. The first of these is based on the principle of Universal Design (American) or Design for All (European) [7], [11]. This principle lays down generic guidelines for designing mainstream products and services, which will accommodate a 'broader average' of users. This means designing mainstream products and services that meet the needs of most users including many of those with disabilities or older people. However, designing for the broader average will not always accommodate the needs of those with severe disabilities or very particular user requirements. Access to technologies for such users can best be achieved by the second means - designing special products and services or in many cases adapting, or interfacing, existing products or technologies to meet the user's specific requirements. The technology based on universal design and on special products and services for older people and persons with disabilities is called 'assistive technology' (AT).

3.8 COST 219ter

The COST Programme [8] has been working on European CO-operation in the field of Scientific and Technical Research. Within COST, Action 219ter [9] has been working since it was established in 2002 in the field of Accessibility for all to services and terminals for next generation networks. The Action not only focuses on telecommunications and tele-informatics, but also in the level of accessibility of different types of contents, such as sophisticated multimedia contents, etc.

3.9 EDeAN (European Design for All e-Accessibility Network)

Created in accordance with one of the specific goals of the e-Europe 2002-2005 Action Plan, EdeAN [12] is mainly engaged in raising the profile of Design for All and emphasizing its importance in achieving greater e- Accessibility. EDeAN consists of a number of Special Interest Groups with the objectives to engange in benchmarking, standardization and proactive assessment in the filed of accessibility.

3.10 OASIS Open Document Format (ODF) Standard

On May 1st, 2005 the Open Document Format for Office Applications (OpenDocument) became an OASIS standard [16]. The OpenDocument is based on XML and performs all functionalities that we need in our office documents. Being XML-based, these documents might be used by other XML-driven applications, like DAISY or NIMAS.

3.11 OASIS Darwin Information Typing Architecture (DITA) Standard

The Darwin Information Typing Architecture [15] became a formal OASIS standard in May 2005. The DITA specification defines not only a set of documents types for authoring and organizing topic-oriented information, but also a set of mechanisms for combining and extending document types. DITA is a working and widely adopted example of content re-use based on XML.

3.12 ISO/IEC JTC1/SC29 (MPEG-4/MPEG-7/MPEG-21)

The MPEG standards are revised and approved under the supervision of SC29 [13], the Committee that deals with the standardization of the coding of audio, picture, multimedia and hypermedia information within the Joint Technical Committee 1

(JTC1) of ISO (International Organization for Standardization) and IEC (International Electrotechnical Commission).

MPEG-4 is a direct consequence of the need of streaming high-quality multimedia through the World Wide Web. Now, MPEG-4 is much more than that and is able to use low bandwidths to transport video and audio to a number of different devices, from Digital TV to a mobile phone.

MPEG-7 is a standard for description of multimedia content. It will be used for indexing, cataloguing, advanced search tools, program selection, smart reasoning about content, etc. It allows the management of high amounts of content locally stored, online and in broadcasts.

MPEG-21 is an emerging standard with the goal of describing a “big picture” of how different elements work together to build and infrastructure for the delivery and consumption of multimedia content. It includes a universal declaration of multimedia content, a language facilitating the dynamic adaptation of content to delivery network and consumption devices, and various tools for making DRM more interoperable. MPEG-21 is about managing content and access to content.

3.13 ISO/IEC JTC 1 Special Working Group on Accessibility

As identified in its long-term business plan and to be responsive to international, regional, national, and end user requirements in the area of accessibility, JTC 1 has established a Special Working Group on Accessibility [14] in order to (a) determine an approach, and implement, the gathering of user requirements, being mindful of the varied and unique opportunities (direct participation of user organizations, workshops, liaisons); (b) gather and publish an inventory of all known accessibility standards efforts; (c) identify areas/technologies where voluntary standards are not being addressed and suggest an appropriate body to consider the new work.

3.14 Accessibility Properties for Learning Resources (APLR)

It is desirable to meet the needs of encoding Accessibility Meta-data for Learning Objects in Europe that APLR [20] harmonize and integrate with an ongoing work in IMS [19], Dublin Core Accessibility, ISO/IEC JTC1/SC36 [22], W3C WCAG (and XHTML), W3 Device Independence Working Group, an ongoing revision of IMS Content Packaging likely to include material suggesting how accessibility-related alternatives can be packaged and other Meta-data efforts including, not detailed here but highly significant, groups working on interoperability of content repositories. Achieving harmonization across all these groups is a very difficult task. A core set of these activities for which progress on harmonization may be more immediately obtainable would include the IMS work, Dublin Core and SC36.

4 Conclusion

There are several standards and initiatives addressing accessibility and it clear at least in the publishing area that structured information is the first big step towards high-quality accessible information. A document whose internal structure can be defined and its elements isolated and classified, without losing sight of the overall structure of the document, is a document that can be navigated. Most adaptive technology allows the user to access a document, and to read it following the "outer" structure of the original. If that structure is left to a range of visual cues, like bold capital letters for the title of a chapter and bold italics for the heading of a subchapter, the adaptive device will surely flatten that visual structure, leaving a document with no structure at all. But if the same document has also an "inner" structure that makes possible for the adaptive device to distinguish between a paragraph and a footnote, between a chapter and a subchapter, then the level of accessibility of the whole document will be greatly enhanced, allowing the user to move through it in the same way those without disabilities do when looking at the printed document, following the same "logic".

In an ideal world, any document made available in electronic format should contain that inside structure that benefits to everyone. Highly-structured documents are becoming more and more popular due to reasons that very seldom have to do with making it accessible to persons with disabilities. Some of the largest publishers are converting their old electronic texts into full XML documents so that it will allow them to look for certain bits of text that they can reuse in further editions, as well as to help them to avoid double-production of the same text. Whatever the reasons behind those decisions are, the use of highly-structured information is of great benefit to anybody accessing them for whatever purpose.

Acknowledgement

This paper acknowledges the European funded EUAIN project and its consortium for contributing and active supporting the DPA workshop, the members of the workshop for delivering substantial input and shaping the work, NEN for being the secretariat of the workshop and CEN/ISSS for hosting and providing the infrastructure for this to happen.

References

1. EUAIN project (European Accessibility Information Network), more info at http://www.euain.org
2. CEN/ISSS WS/DPA Business plan, available at http://www.cenorm.be/cenorm/businessdomains/isss/activity/dpan0010approvedbusinesplan1.doc
3. Martinez Calvo, FM., (ed) (2005) Standards for Accessible Information Processing, EUAIN Consortium, FNB Amsterdam
4. AIIM PDF Universal Access Working Group, http://www.aiim.org/standards.asp?ID=27861
5. CEN/ISSS - European Committee for Standardization http://www.cenorm.be/cenorm/aboutus/index.asp
6. CEN/ISSS MMI-DC http://www.cenorm.be/cenorm/businessdomains/businessdomains/isss/activity/wsmmi.asp
7. CEN/ISSS DfA Workshop-CEN/ISSS Design for All and Assistive Technologies in ICT, at http://www.tiresias.org/guidelines/cenisss/

8. COST - European CO-operation in the field of Scientific and Technical Research http://cost.cordis.lu/src/home.cfm
9. COSTter219 - Accessibility for All to Services and Terminals for Next Generation Networks http://www.tiresias.org/cost219ter/mtgs.htm
10. DAISY - Digital Accessible Information System, at http://www.daisy.org
11. Design for All - http://www.designforall.org
12. EdeAN - European Design for All e-Accessibility Network http://www.edean.org/index.asp?auto-redirect=true&accept-initial-profile=standard
13. ISO/IEC JTC1/SC29 - Subcommittee 29 of the Joint Technical Committee 1, http://www.itscj.ipsj.or.jp/sc29/open/29view/29n4036c.htm
14. ISO/IEC Special Working Group on Accessibility - http://www.incits.org/SWG-A/documents/swga_027.pdf
15. OASIS DITA standard - OASIS Darwin Information Typing Architecture Standard http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=dita
16. OASIS ODF standard - OASIS Open Document Format Standard, http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=office
17. WCAG 1.0 - Web Content Accessibility Guidelines 1.0 http://www.w3.org/TR/WAI-WEBCONTENT/
18. WCAG 2.0 - Web Content Accessibility Guidelines 2.0, http://www.w3.org/TR/WCAG20/
19. IMS Global Learning Consortium, http://www.imsglobal.org/content/packaging/
20. APLR: Accessibility Properties for Learning Resources, http://axelrod.cms.shu.ac.uk/aplr/
21. WAI: Web Accessibility Initiative, http://www.w3.org/WAI/
22. ISO/IEC JTC1/SC36, Subcommittee 36 of the Joint Technical Committee 1, http://jtc1sc36.org/
23. WCAG 2.0, Web Content Accessibility Guidelines 2.0, http://www.w3.org/TR/WCAG20/

Expressing Emotions Using Animated Text Captions

Raisa Rashid, Jonathan Aitken, and Deborah I. Fels

Ryerson University, 350 Victoria St., Toronto, Canada M5B 2K3
dfels@ryerson.ca

Abstract. Due to limitations of conventional text-based closed captions, expressions of paralinguistic and emotive information contained in the dialogue of television and film content are often missing. We present a framework for enhancing captions that uses animation and a set of standard properties to express five basic emotions. Using an action research method, the framework was developed from a designer's interpretation and rendering of animated text captions for two content examples.

1 Introduction

While many factors contribute to the moulding and propagation of North American culture, television and film are, arguably, the most prominent and pervasive media, influencing, informing and entertaining the audiences they attract. Whether this influence is strong or weak, these two media give individuals access to information that is readily and easily available. The majority of North Americans have little difficulty accessing the medium's message as it is presented using audio and visual content. People who are deaf or hard of hearing, however, experience limited access to these media and, because of this, are unable to fully participate in the common cultural experiences that television and film offer. Technologies that facilitate the transmission of audio content to people who are deaf/hard of hearing need to be more adequately developed to allow these individuals to enjoy television and film. Current North American closed captioning does provide some access to the sound information. However, it is limited and does not sufficiently meet all needs of this community.

Closed captioning for analogue television (EIA-608 standard) is the standard mechanism used to provide access to television and film (Field, 2000). The current form of closed captioning uses simple text-based format, with a modified character set built into the television decoder chip. It consists of a single mono-spaced font, a single font size, a short set of text colors, and a black background (VITAC, 2003). Initially, the system was limited simply to white uppercase letters. Since this original specification, several options have been added, permitting use of mixed case letters, a small set of text colors, along with a few special characters (e.g., music notes). These additions, however, are frequently absent, since users have become habituated to and prefer uppercase captioning and a white font. It is only recently that caption writers have developed an interest in using mixed-case words.

Teletext has been prevalent in Europe for decades, offering text-based information such as the news, weather in addition to subtitling. The teletext service allows viewers to access "pages" of information using their remote control, and specific "pages"

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 24 – 31, 2006.

are dedicated for captioning purposes (NCAM, n.d.). Unlike line 21 closed captioning, the current standard in North America, teletext subtitling offers more flexibility and features such as color and animated icons due to a higher data transmission rate (NCAM, n.d.). In fact, viewers of teletext are quite accustomed to seeing these additional features used to convey information beyond dialogue, such as speaker identification through the use of different colored text.

Jordan, Albright, Branner, & Sullivan (2003) suggested that although audiences were generally satisfied with captioning quality for analogue television, missing words, spelling errors and captions moving too quickly caused dissatisfaction. Instead of just using text to provide all sound descriptions, it may be possible to use symbols, icons, animation or moving text as ways to express sound information. These alternative techniques could be used to decrease the amount of text-based captioning required and, consequently, reduce the rate at which captions move. These alternatives could also be used to capture some sound information such as music and sound effects that cannot easily be described using text (often because there is no time or space in text captions for additional descriptions).

To understand the full message being communicated with existing closed-captioning models, the deaf or hard of hearing caption viewer must rely on visual-only cues such as body language and gesture, and combine this information with words and short text descriptors shown within the captioning. Mehrabian, (1968) suggests that successful human-to-human communication consists of words, non-verbal linguistic modifiers (paralanguage), and visual cues or gestures with a majority of the meaning of that communication being derived from paralanguage, and visual cues and gestures. Cruttenden (1997) also discusses the importance of prosodic or paralinguistic features of length, loudness, and pitch in imparting meaning to spoken language. It is therefore important to ensure that paralinguistic information is captured in captions or teletext so that more of the meaning of the text can be understood.

Fels, Lee, Branji, & Hornburg (2005) have begun to investigate the use of alternative displays for expressing emotion, music and sound effects for digital television. Initial results are very positive as viewers who are deaf and hard of hearing seem to have a different and more satisfactory experience of the captioned video content. In addition, these researchers have developed software technologies that allow for emotive elements to be indicated in a script or text and then automatically rendered into graphical captions. However, they have only explored the use of graphics and icons for captioning and not other forms such as animation and animated text (or kinetic typography).

In this paper, we present a framework that relates the properties of kinetic typography to the emotions present in speech. This framework was developed using a creative design approach with two example pieces of content and was informed by a model of basic emotions.

1.2 Kinetic Typography

Kinetic typography refers to animation and movement of type in order to add meaning, tone of voice, or emotion to written words. The goal is to approximate varying levels of meaning inherent in spoken word as opposed to written text. A person's voice, for example, is affected by many variables such as tone, rhythm, pitch, and volume--all of which affect meaning. Animated text can also convey similar expression of content through motion.

Existing research on kinetic typography focuses on three key areas: studies examining its effect on experienced emotion; research linking concept of "voice" to animated text; and development of software to simplify its creation. Current research into kinetic typography and its effect on experienced emotion is relatively recent. Using an accepted mood rating scale (Plutchik, 1980; and R. Plutchik & Kelleman, 1989) and video taped facial expressions, Stone, Alenquer, Borisch (2004) compared emotional responses to static and animated words. Their results showed a small increase in emotional response to moving words. However, in this study only words were presented to subjects with no context. It is difficult to assess emotional responses out of context because as Mehrabian (1968) suggests paralinguistic information is an important element in communication systems.

Wang, Prendinger, Igarashi, (2004) used real-time galvanic skin response measures, which detects arousal using skin conductivity. Study results showed that the emotional response of the user was heightened for kinetic type compared with static type.

Research is also being carried out on how animated text can be used to convey a sense of "voice" to the written word. Forlizzi, Lee and Hudson (2003) present an overview of research into the relationship between voice, emotion and kinetic typography. Generally, they explore the success of kinetic typography's capacity to convey tone of voice; examining paralinguistic and prosodic, or linguistic, features. Forlizzi, et al (2003) found that it is difficult to portray paralinguistic features such as a "husky" vocal quality, but found more success with conveying linguistic features such as pitch, loudness and tempo.

1.3 Primitive Emotions

In our research it was necessary to use a simplified model of emotion to characterize the emotional elements present in television and film dialogue. Psychological models of emotion such as those proposed by (R. Plutchik, 1980) and (Ekman, 1999) suggest that all emotions can be reduced to a set of five to eight primitive emotions. These are anger, sadness, happiness, fear, surprise, disgust, anticipation and acceptance, although there is considerable variation of the number and labels of these primitive emotions (Acton, 1998). In more than 50% of the studies cited by Acton (1998), anger, sadness, happiness and fear are common denominators. These emotional categories thus formed our simplified model of emotions for captioning. We also found that a disgust category (suggested as a primitive emotion by Plutchik, 1980) was required for our emotion model as it appeared many times in the example video content, and required a unique animation style.

2 Creative Process

Because the literature offered little guidance on how to associate animation with specific emotions, we decided to use a more design-oriented and intuitive approach that involved an experienced graphic designer and creative team. This action research approach consisted of generating a set of alternatives, selecting one concept based on group consensus, application of that concept to captioning content, refinement and then final production. While there were natural limitations to this creative process, it seemed the most useful and productive method at this initial stage. The following account is a brief description of the process undertaken by the creative team

Although subjective in nature, there did seem to be some movements that suggested a specific emotion to the creative team. We found that these movements loosely paralleled patterns of speech intonation. Juslin & Laukka (2003) found that rising vocal contours were associated with anger, fear and joy in most studies they surveyed. Motions that echoed this pattern seemed to suggest similar emotions. For example, type that increased quickly in size and "vibrated" rapidly suggested fear. Slower movement using the same style seemed to indicate less fear. These apparent shared perceptions formed a good starting point for the generation of ideas. At this stage, we focused on generating many alternative and specific example concepts. We began with rough sketches and from these produced many simple "test" animations. This period of ideation lasted several weeks.

Next, we examined the work more critically and distilled the concepts to the strongest directions. We then applied them to video samples and presented this work to the full research team consisting of three individuals with some experience in closed captioning and audiences who are deaf or hard of hearing. Here, for the first time, we solicited critical input from people who had little experience and exposure to animated text and its intended meaning. Their reactions were essential in informing the directions to take, and in judging the potential effectiveness of the various alternatives.

Once agreement was reached on which alternatives to adopt, they were then applied to the text captions of two one-minute examples of video content. As we progressed through this task, it became evident that some emotions did not lend themselves easily to animation. For example, the motion representing sadness was not consistently rendered, nor was it accurately interpreted by the group. In one instance, for example, motion was used to move type "up" on the page, to complement dialogue which said "we're all going to die!" said with an increasing pitch in the voice. The motion was meant to convey that sense of "dying" but was too tenuous a link for understanding. To correct it, we discussed linguistic cues that suggested sadness and settled on falling intonation. This then became downward motion. These and other ideas were then applied and the animation was refined until there was agreement on the final animations. This refinement stage continued for several more weeks, until the final animations were complete.

We believe that this creative process has resulted in animations that communicate several basic emotions; however formal user testing is still required to determine whether an audience who is deaf and hard of hearing can consistently interpret the emotive meaning of these animated text styles.

3 Framework of Emotive Captions

Once two content pieces were captioned using the expert's approach, we attempted to create a framework relating animation properties to the set of eight basic emotions as described by Plutchik (1980). However, three emotions surprise, anticipation and acceptance did not produce unique animation properties but existed within the other five. As a result, only five emotions, anger, fear, sadness, happiness and disgust comprised the framework.

The animation properties consisted of a set of standard kinetic text properties. These are: 1) text size; 2) horizontal and vertical position on screen; 3) text opacity or how see-through the text is; 4) how fast the animation moves/appears/disappears (speed/rate); 5) how long it stays on the screen (duration); and 6) a vibration effect. Vibration is defined as the repeated and rapid cycling of minor variations in size of letters and appears as "shaking" of the letters.

All of the properties are defined for three stages of onscreen appearance; 1) onset or when the text first appears or begins to animate; 2) "the effect" occurring between the offset and onset stages; and 3) offset or when the animated text stops being animated or disappears from the screen. In addition, the intensity of the emotion can affect all of the properties. For example, high intensity anger has a faster appearance on the screen and grows to a larger text size in the effect stage than text that shows a

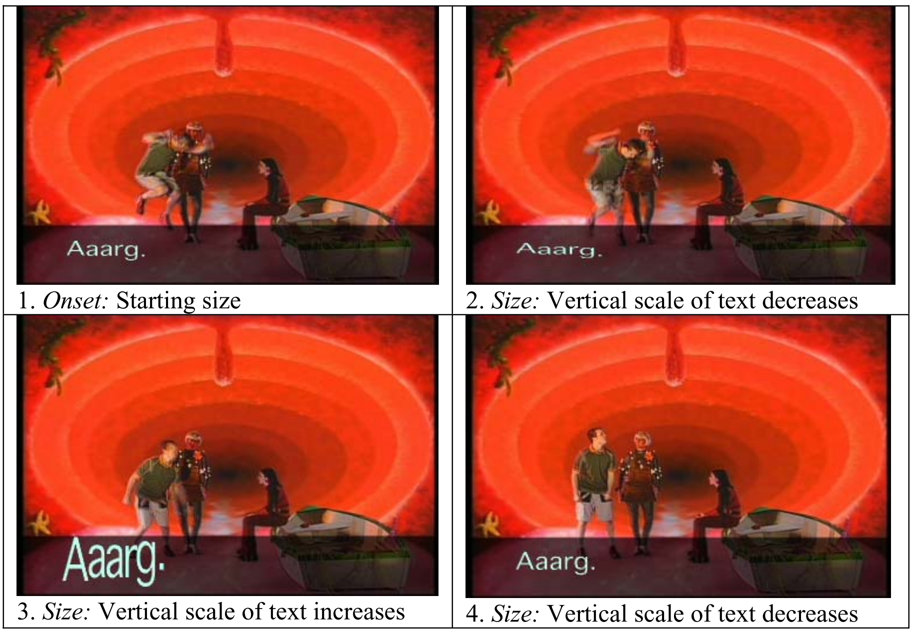
Table 1. Summary of the relevant animation properties for anger, fear, sadness, happiness and disgust

Emotion	Relevant Properties (Effect)	Intensity of Effect
Fear	*Size*: Repeated and rapid expansion and contraction of the animated text. *Rate*: Expansion and contraction occur at constant rate. *Vibration:* Constant throughout the effect.	*Low*: Size of animated text is the same as non-animated text at onset. Low level vibration. *High:* Size of animated text is larger than non-animated text at onset. High level of vibration.
Anger	*Size:* Abrupt expansion and contraction of the animated text (through one cycle). *Vibration:* Occurs at largest size in cycle. *Onset:* Fast onset *Duration:* Pause at largest size in cycle and vibration occurs here.	*Low*: Size of effect is smaller. Slower onset. *High:* Size of effect is larger. Faster onset.
Sad	*Size:* Vertical scale of text decreases *Position:* Downward vertical movement from baseline. *Opacity:* Decrease in opacity *Offset:* Slow offset.	*Low:* faster offset *High:* slower offset.
Happy	*Size*: Vertical scale of text increases *Position:* Upward vertical movement from below baseline. Follows a curve (fountain-like).	*Low:* Slower onset. Offset text size is smaller. *High*: Faster onset. Offset text size is larger.

lower intensity anger condition. The text will also seem considerably larger than the surrounding text. Many of these properties are defined relative to the original size of the non-animated text and they would be identified as a default value by the captioning standard that is applied.

Table 1 summarizes all of property descriptions for each primitive emotion. Where an animation property is not defined, a default value would be used. For example, where onset speed is not specifically defined, this would be the default onset speed defined by the captioning software. The framework also allows for other emotions to be added and the related animation properties defined accordingly.

Figures 1 and 2 illustrate these concepts for a high intensity anger caption and a low intensity fear caption with the most salient properties identified. The animations would appear during the time that the captions were present in the video material.



1. *Onset:* Starting size

2. *Size:* Vertical scale of text decreases

3. *Size:* Vertical scale of text increases

4. *Size:* Vertical scale of text decreases

Fig. 1. Example of high intensity anger over four frames of play. The text size shrinks for its default onset size, expands and shrinks again defining one cycle. Vibration occurs when the text is at it largest size.

Animations created using this framework could be applied to words, sentences, phrases or even single letters (grain size) in a caption to elicit the desired emotive effect. In our examples, we demonstrated that the intended emotive effect could be consistently applied for different typefaces (e.g., Frutiger, Century Schoolbook).

However, whether an audience gains a better understanding of the emotive context of the content from animated captions remains to be evaluated. In addition, important questions relating to the practical application of this approach are: 1) whether there is a relationship between the intensity of the emotion, the grain size to which the animation should be applied to reflect that emotion, and the quantity of dialogue containingthat emotion; and 2) how much animation is enough to bring improved understanding to an audience but not to cause distractions due to over animated text.

1. *Size:* Expansion of text. *Vibration:* Constant	2. *Size:* Contraction of text. *Vibration:* Constant

Fig. 2. Example of low intensity fear. Initial text size is default size. Text size then expands and contracts rapidly for the entire duration that the text is on the screen.

4 Conclusion

Although these concepts remain to be evaluated with users, we have shown that it is possible to animate captions in a way that represents an designer's interpretation of those emotions. Based on the application of this approach to two different pieces of video content, we developed a framework that associates basic or primitive emotions with the standard properties of animation providing a consistent method for applying animation to text captions. Next steps in this process involve user testing with existing examples, and examples from different genres of content to refine the framework and then creating animation tools to implement the framework with standard video editing tools, caption editors and online tools such as Flash.

Acknowledgements

Funding for this project was partially provided by NSERC, grant # 184220-05 and Ryerson University. The video content was generously provided by Marblemedia Inc. Advice was provided by Jim Hardman from the Canadian Hearing Society.

References

1. Acton, G. S. (1998, April 1998). *Basic Emotions*. Retrieved Feb. 25, 2005, from http://www.personalityresearch.org/basicemotions.html
2. Ekman, P. (Ed.). (1999). *Chapter 3:Basic Emotions*. Sussex: John Wiley & Sons.

3. Fels, D. I., Lee, D. G., Branji, C., & Hornburg, M. (2005). *Emotive captioning and access to television.* Paper presented at the ACMIS 2005, Omaha
4. Field, G. (2000). *Implementing DTV Closed Captions and Video Description*. Retrieved Feb. 28, 2005, from http://ncam.wgbh.org/dtv/overview/nab2000paper.html
5. Forlizzi, J., Lee, J., & Hudson, S. (2003). The kinedit system: Affective messages using dynamis texts. *Chi 2003: New Horizons, 5*(1) 377-384
6. Jordan, A. B., Albright, A., Branner, A., & Sullivan, J. (2003). *The state of closed captioning services in the United States*: National Captioning Institute Foundation.
7. Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin., 129*(5), 770-814
8. Mehrabian, A. (1968). Communication without words. *Psychology Today, 2*(4), 53-56.
9. Plutchik, R. (1980). *Emotion, a psychoevolutionary synthesis*. New York: Harper & Row.
10. Plutchik, R., & Kelleman, H. (1989). *The measurement of emotions*. New York: Academic Press
11. Stone, B., Alenquer, D., & Borisch, J. (2004). Type, motion and emotion: A visual amplification of meaning. In D. McDonagh, P. Hekkert, j. van Erp and D. Gyi (Eds.), *Design and emotion the experience of everyday things*. New York: Taylor & Francis.
12. VITAC. (2003, April 2003). *History of DTV Captioning*. Retrieved Feb. 28, 2005, from http://www.vitac.com/technical/dtvhistory.htm
13. Wang, H., Prendinger, H., & Igarashi, T. (2004). Communicating emotions in online chat using physiological sensors and animated text. *Extended abstracts of the 2004 conference on Human factors and computing systems Late breaking result papers.* 1171-1174

Schulbuch Barrierefrei (Accessible School Books) – Co-operation Between Publishers and Service Providers in Austria

Klaus Miesenberger and Reinhard Ruemer

University of Linz, Institute Integriert Studieren
Altenbergerstr. 69, 4040 Linz, Austria
{klaus.miesenberger, reinhard.ruemer}@jku.at

Abstract. This presentation outlines the co-operation between Austrian school book publishers and service providers for people with special needs which aims at making books available in electronic format. A project has been working on a) a minimum set of structural elements to be incorporated by publishers, b) know-how and handouts for publishers how to implement structured design using standard desktop publishing (DTP) systems, c) examples of new books and redesign existing books, d) training materials, workshops and seminars to transfer the know-how to other publishers and design agencies, e) a general agreement which gives the right of transferring books in electronic format to students with disabilities, f) a Document Rights Management System, g) a workflow for the co-operation between schools/teachers, service providers, publishers and the ministry.[1]

1 State of the Art

1.1 Accessibility and Structured Document Design

Computers have changed the way we write, distribute, store and handle information. The Gutenberg Age in publishing which was exclusively based on typeset and printed documents seem to be over. Computer Supported Publishing and Electronic Publishing are state of the art and using Information and Communication Technology (ICT) offers new possibilities to give access to information to people disabilities using Assistive Technologies (AT).

The Gutenberg Age led to and excluding tendency as documents had to be reproduced in alternative formats. Only a small part of these documents was made available and the specialty of access and handling of these alternative formats can be seen as one of the fundamental reasons for exclusion [1, 2].

ICT and AT and therefore electronic documents are of a basic inclusive tendency. Only in the moment of access the media and handling qualities are put in place at the Man-Machine-Interface including AT. Presentation and handling become independent from the content. People with disabilities can use the same document as all other people – an inclusive tendency.

---

1 A second part of the project focused on accessible design of online materials, additions to standard printed documents. We do not report on this part of the project here.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 32 – 39, 2006.

This potential for people with disabilities is based on structured document design, the use of mark-up and descriptive languages which forms this independence of content, media and handling [2,3,4]. TeX [5] has been a first application revolutionizing the publishing sector and opening possibilities to access electronic documents or to automate the production of alternative formats. Other important examples of early approaches (in the 60ies and 70ies) were done using SGML [6] or EMACS [7] which provoked a series accessibility work [3,7,8] in the 80ies and 90ies. The following steps towards hypermedia and the World Wide Web [e.g., HTML, XML, XHTML] more and more see accessibility as an integral part. Accessibility converges with mainstream requirements of a more flexible and adaptable access to information. The DAISY consortium [9] and the World Wide Web Accessibility Initiative WAI [10] in particular, but also other approaches like e.g. TEI [11] or the International Digital Publishing Forum [12] (formerly the Open eBook Forum) work towards standards offering an open, flexible and therefore accessible access to documents.

Nevertheless the publishing sector does not support these standards and therefore the accessibility potential today. But due to the need of using their sources for different publishing purposes in the future (e.g., print, online, CD/DVD) publishers start to take up these possibilities of “Multi Channel Publishing” [13] which by their nature converge in many aspects with accessibility requirements. The DAISY consortium and nowadays the EUAIN project [14] are examples of initiative to mainstream accessibility into the publishing process.

1.2 Situation in Austria

Since several years the school book production in Austria tries to use these possibilities to make books accessible to students with disabilities and to put an according co-operation with publishers in place.

In Austria, the Federal Ministry for Social Affaires and Generations is providing educational materials like schoolbooks and other materials for primary and secondary education for free. Blind and visually handicapped students – and hopefully soon in the future other print disabled students - can order books in accessible formats. Accessible formats are

- Braille books
- Enlarged copies of the original (color, up to A3 format)

Publishers, till this project, did not agree on handing over and distributing digital copies of books. The development of Braille books starts from printed books with scanning, OCR or, when lots of graphics and/or formal structures like math are used, with typing. In this process structure was added to the book, headings were defined and lists and other structural elements were assigned to the text. This has been a very time consuming process.

A more efficient alternative would be

- to use the electronic source documents from the publishers to create an accessible version for students with special needs and
- to hand over electronic versions of books to the students which they then themselves in co-operation with their teachers could adapt to their specific needs.

This would also be the basis for extending the user group to other groups of students with disability. The work would orient much more to a design for all approach which might in the long run also support the publishers in using their document sources for different purposes (e.g. print, online, DVD publishing). Due to this additional focus the project establishing the co-operation with publishers was called "Multi Channel Publishing" [13] what outlines the convergence of interests and shows that producing accessible formats is much more an integral part of the general publishing process.

An introductory study conducted by the Institute Integriert Studieren in 2003 on PDF source documents from the publishers showed that these PDF's are only of very limited or no use for

- students and
- the production process of accessible versions

due to the lack of structural definition. The quality of the PDF, depends on the quality of the structural design of the source document. Only if the original document uses a good structure it is transferred to PDF or any other format in the conversion process. The quality of the documents was that bad that most of the time text was not exported in the right order. The study showed that structured document design is not practiced at publishers' or design agencies' site as it simply is not needed at the moment. Instead of using the features to structure the content authors or typesetters just use visual styles.

These results motivated to start a project which addressed the issues listed in the summary. Publishers are interested to take part as a) the new anti discrimination legislation [20] will ask for accessibility of school books and b) they experience general problems in the publishing process when they want to use sources for different publishing purposes (e.g. print, online, CD, audio/multimedia). This convergence of interests led to a strong partnership for the project named:

2 The "Multi Channel Publishing" Project [13]

Five publishers take part in the project. Each of them is responsible for designing or redesigning one of their books based on a predefined set of structural elements. This basic structural design defined in the project guarantees that the electronic version of the book can be used for the production of alternative formats. An analysis of the publishing process at publishers' site showed that service providers can only start from the final print ready version as the content, which is approved by public authorities, changes till this point. This final version today is most of the time a PDF generated from a DTP Tool (e.g. Adobe InDesign [15] or QuarkXPress [16]). Due to this, if electronic sources should be usable for services providers, structured design has to be implemented into the DTP work.

2.1 Definition of Structural Elements for Electronic Versions of Books

To be able to collect the data of the source document and convert it into a XML File, we used the element Set of the TEI-Standard [11], in particular the TEI Lite DTD.

The TEI's Guidelines for Electronic Text Encoding and Interchange were first published in April 1994. This set of meta data is widely know by publisher and guarantees compatibility or convertibility to other definitions in use like DAISY [9]. Using TEI keeps the process close to the upcoming XML database schemes which publishers might use in the future using database structures for processing their documents. The TEI Lite DTD still consists of over 120 Elements for the tagging of books, most of them important for librarians. To simplify the work for all participating parties, a subset of those elements was selected. This subset consists of structural elements which are of general importance for structured document design and automatic content processing. This subset does not ask for special knowledge of accessible versions but can be seen as the basis for structured document design in general. Using this subset guarantees that the sources (or PDFs) can be used as a starting point for the production of accessible versions. In generals this sub-set of the TEI Lite DTD comprises structural meta data elements for:

- Headings
- Divisions / Subdivisions
- Images
- Tables
- Notes
- Page breaks
- References

It also comprises administrative meta data elements (e.g. Edition, Year of Publishing, Author(s), Publisher,…). The experience in the project showed, that this DTD Subset is sufficient to structure the content of the schoolbooks. Publishers after a short training were able to do the work by themselves. This subset also proved to be in accordance with new publishing systems based on XML databases.

2.2 Authoring Tools

After the definition of the XML DTD, knowledge was developed how the authoring tools could support efficiently marking up documents in the right way during the layout process. Further on routines for exporting the defined structure and layout data into XML. The two most widely used authoring tools were examined in detail:

Adobe InDesign. InDesign from Adobe Inc. is a desktop publishing application (DTP) which can work with XML files. It is possible to import XML into InDesign and then prepare the document for output e.g. printed book. This feature is an important step toward multi-channel and cross-media publishing. Tests with Adobe InDesign CS 2 showed that it is possible to tag the text of the layout document. Further investigations are done to efficiently map layout to the structure. InDesign supports the mapping of text-formats to XML-Tags but the structure had to be added afterwards. The mapping feature can be used, if the text is in a proper layout. Otherwise the user has to mark the specific text area (e.g. one chapter) and then to assert the XML tag to the text.

QuarkXPress. QuarkXPress is another desktop publishing application (DTP) produced by Quark Inc. With QuarkXPress users can import and export XML

Documents. With Quark Digital Media Server content can be stored in a central database. It then can be used in multiple forms according to the principles of multichannel publishing.

Quark XTensions software, which are plug-ins, can automate functions and eliminate repetitive steps with palettes, commands, tools, and menus. Tests with QuarkXPress 6.5 Passport (international Edition) showed that QuarkXPress was not able to import the TEI-DTD. To tag the text of the book, a new, flat DTD had to be written. With the new DTD the mapping from layout formats to XML tags was possible. The content then is exported into a XML file. This is the basic version for the accessibility work.

2.3 Example Books

The post-processing tasks are necessary, because, as mentioned before, the exported files in some cases have no structure and there are also parts of some books that could not be exported (e.g. graphics, made in the authoring systems). The post-processing tasks were:

- Adding Structure to the XML
- Revise elements, that were not exported properly
- Describe Images

The result after the completion of the work is a valid XML version of the book.

The next step is to convert the XML via style sheets into the target format. The style sheets for the conversion are freely available on the internet [17]. They allow to convert the XML file into a HTML file with one/multiple pages and also to convert the XML file into a PDF file.

2.4 Training Materials, Seminars and Workshops

Training materials have been developed which are now used in workshops and seminars to transfer the knowledge to as many publishers as well as design agencies as possible.

2.5 DRM-System

DRMs are important to get access to documents owned by publishers [18,19]. To make sure that the books are not used outside the designated user group a DRM System was customized. The system consists of a secure-reader-software and a USB dongle, which acts as the key. Every student gets a key and the software. The key has a code, which allows the student to read the book, if the key is plugged into the computer. This system has the advantage, that the user is not bound to one specific computer or peace of hardware. He can read the book for example at school but also in a learning group or at home. How the students get their books and a detailed workflow between publishers and the service providers is described in the next paragraph.

2.6 Workflow

To start the process, a teacher of a student with special needs orders a book in an accessible format. If the schoolbook service provider does already have the book in its stock, it will be provided directly to the student. Otherwise, the service provider asks the publisher for the electronic version of the book. The publisher sends his TEI-XML file to the service Provider. The service Provider produces the accessible version of the book. Printed (Braille/enlarged) copies are sent by standard mail. If an electronic document is ordered, the service provider encodes the files with the DRM system using the data from the USB dongle of the student. The book is placed on a server, where the student can download the book. When the student has the reader software installed and the dongle plugged in, he can open the book and read it.

2.7 Agreement Between Publisher and Service Provider

To ensure that the process works efficiently, an agreement between publishers and service providers has been worked out. The core articles of the agreement are:

- The publishers provide their electronic source documents
- It must be ensured, that the books are only given to persons with special need
- A DRM system must be used therefore
- It must be a "closed" system with registered users

The agreement will be signed by every publisher and service provider. If a service provider needs a book from a publisher he can ask for it under the condition of the framework agreement.

3 Conclusion

The most important result of the project is the fact that handing over digital copies of print published documents will take place in the future.

The project showed that it is technically feasible to create XML versions of books by using the print ready version of a document. The experience also showed that the quality of the XML after just using the functions provided by the authoring tools is not good enough. A lot of work has to be done afterwards by cleaning and revising the XML document. The persons who are performing this work will have to have some basic XML skills. It will also be a challenge to convince the publishers to create documents that can be exported into XML without a lot of additional effort. In some areas at the moment there are only limited possibilities to sources from publishers, especially in areas, where books consist mainly of pictures, graphics and other visual content. Another challenge is the integration of non-text content like mathematical or chemical expressions.

Working with source files of publishers also showed that the quality differs a lot depending on the awareness and skills of the staff in well structuring and formatting the source files for multi channel publishing. It might well be that it is more efficient in the production of alternative formats to start with scanning and OCR.

The project made obvious that all publishers pass their layout data to the print office by using PDF. An important task in the future will be, to allow authoring systems to create PDF files that are either accessible or allow a conversion back into a useful format.

In any case these are only first, but important steps towards multi channel publishing. More work is needed for a more efficient production of different versions of one source document.

Acknowledgment

This project has been supported by the Austrian Federal Ministry for Social Affaires and Generations.

References

1. Miesenberger, K.: Applied Computer Science for Blind and Visually Handicapped People - Information Technology Enabling Integration at University, Universität Linz, 2001
2. Darzentas, J., Miesenberger, K.: Design for All in Information Technology: a Universal Concern (Keynote), in: Andersen, K., V., Debenham, J., Wagner, R. (eds.): Database and Expert Systems Applications, 16th Internaitonal Conference, DEXA 2005, Copenhagen, Denmark, August 2005, Proceedings, Springer LCNS 3588, Berlin/Heidelberg 2005, pp. 406 – 420
3. Ayres, N., Wesley, T.: Open Document Architecture – An Opportunity for Information Access for the Print Disabled, In: ICCHP, Vienna 1992, pp. 556-561
4. Engelen J. J., Evenepoel F. M., Allemeersch F.: Standardisation of Accessibility: the Tide CAPS and HARMONY Projects, 2nd TIDE Congress, Paris, 26-28 April 1995, The European Context for Assistive Technology, pp. 59-65, IOS Press, Amsterdam, ISBN 90-5199-220-3
5. Knuth, D. E.; Bibby D. R.: The TEXbook. ADDISON-WESLEY, Reading, Mas., 1983
6. OASIS homepage: Standard Generalized Markup Language (SGML), http://xml.coverpages.org/sgml.html, available on January 21, 2006
7. Emacspeak - The Complete Audio Desktop, http://emacspeak.sourceforge.net/, available on January 21, 2006
8. T. V. Raman: AsTeR - towards modality-independent electronic documents, presented, at DAGS. In: http://simon.cs.cornell.edu/Info/People/raman/publications/DAGS-95/paper.html, 1995
9. The Daisy Consortium Web Page, http://www.daisy.org/, available on January 21, 2006
10. The WWW Accessibility Initiative Homepage, http://www.w3.org/WAI/, avaiable on January 21, 2006
11. The Text Encoding Initiatve Home Page, http://www.tei-c.org/, available on January 21, 2006
12. International Digital Publishing Forum (formally Open eBook Forum) homepage, http://www.idpf.org/ available on January 21, 2006
13. Miesenberger, K.: „Multi Channel Publishing“ - für die Nutzung elektronischer Dokumente für Menschen mit Behinderungen, Erarbeitung von Richtlinien, Trainings- und Begleitmaterialien und Beispielen für die Erstellung strukturierter Bücher und Online-Ergänzungsmaterialien, Projektantrag, Universität Linz, 2004
14. European Accessible Information Network homepage, http://www.euain.org/, available on January 21, 2006
15. Adobe InDesign. Homepage. Online: http://www.adobe.com/products/indesign/main.html
16. QuarkXPress. Homepage, http://www.quark.com/products/xpress/, available on January 21, 2006

17. Text Encoding Initiative - XSL stylesheets for TEI XML. Homepage. http://xml.web.cern.ch/XML/www.teic.org/Stylesheets/teixsl.html, available on January 21, 2006
18. Miesenberger K., Burger, F.: Secure Distribution System for Publications in Electronic Form, in: Klaus, J., Auff, E., Kremser, W., Zagler, W., L. (Eds): Interdisciplinary Aspects on Computers Helping People with Special Needs, 5th International Conference, ICCHP ´96, Oldenbourg, Wien, München 1996
19. Bormans, G., Engelen, J., Wesley, T.: SEDODEL: Technical solutions to the copyrigth barriere for widespread electronic information, in: Vollmer, R., Wagner, R.: Computers Helping People with Special Needs, ICCHP 2000, Oldenburg, Wien, 2000
20. Behindertengleichstellungsgesetz – Regierungsvorlage (government bill), 2005, http://www.parlament.gv.at/portal/page?_pageid=908,848700&_dad=portal&_schema=PORTAL, online available on May 13, 2005

Modelling Accessibility Constraints

Terje Gjøsæter, Jan Pettersen Nytun, Andreas Prinz, Mikael Snaprud, and Merete Skjelten Tveit

Agder University College
Grooseveien 36
Grimstad. Norway
{terje.gjosater, jan.p.nytun, andreas.prinz, mikael.snaprud, merete.s.tveit}@hia.no

Abstract. This paper describes the combination of research topics from two projects focusing on web accessibility testing, and on metamodelling and high-level specification languages. This combination leads to a novel approach to accessibility assessment that will improve the understanding of accessibility issues, and explore the potential for generating executable accessibility tests based on accessibility constraint models.

1 Introduction

Access to web content for all is crucial for building an inclusive information society. Information on the web should be accessible to all users, independent of disabilities or choice of web browser. In order to improve accessibility, the EIAO [1] project evaluates web content according to accessibility. This will raise the awareness for accessibility, and hopefully in the long run increase accessibility.

However, the evaluation of web pages is only trustworthy if the measurements are transparent to the users so they can understand how a ranking is established. We see two supplementary ways to achieve transparency: 1) the consistent use of Open Source software, and 2) explicit models of tests.

The SMILE [2] project is concerned with explicit representation of semantic information, such as description of formulae or algorithms. This is done using a metamodelling approach with an explicit description of the semantics of a language using a model of the language.

Building on experiences and results from the EIAO project, we apply the methodology of the SMILE project in the area of the EIAO project. Describing the issues with a metamodel-based approach [3], we will be able to model accessibility aspects directly. This work will result in a platform for experiments with accessibility constraints.

2 The EIAO Project

The EIAO project is designed to develop a European Internet Accessibility Observatory (EIAO). It comprises a set of Web accessibility metrics, an Internet robot for automatic collection of data on Web accessibility and deviations from Web

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 40 – 47, 2006.

accessibility standards, and a data warehouse providing on-line access to measured accessibility data. In this way, the EIAO project will contribute to better e-accessibility for all citizens in Europe by providing a systematic and automatically updated overview of the current compliance with Web accessibility guidelines. EIAO is carried out within the Web Accessibility Benchmarking (WAB) Cluster [18] together with the projects Support-EAM [4] and BenToWeb [5], also co-funded by the European Commission[1]. The cluster consists of 24 partner organisations in Europe.

Web accessibility is on the European agenda through the eEurope action plans 2002/2005 on European and on national level. eGovernment policies in many European countries require compliance with WAI [6] standards, and the private sector increasingly has to deal with Web accessibility. Existing automatic Web site tests are mainly used to evaluate and improve single Web sites. Benchmarking of, for example, all public sector Web sites or large collections of Web sites are not available. The objective of the European Internet Accessibility Observatory EIAO is to develop a framework for an automated Web accessibility assessment. The EIAO assessment system can provide valuable input to European benchmarking and can effectively support Web accessibility analysis and policies. Together with the EU projects BentoWeb and Support-EAM, the EIAO project forms a cluster that develops a Unified Web Evaluation Methodology (UWEM) [15].

3 The SMILE Project

The SMILE project accepts the current challenge of semantic information processing. Providing a methodology combined with a framework, SMILE allows us to enrich data with semantic interpretation (context knowledge); this is a key technology for providing tools for IT-based information processing. Using an extended project-specific metamodel approach, SMILE supports turning low-level information into higher-level information in new contexts by generating tools that produce and process this higherlevel information.

The capability of the SMILE methodology and of the framework will be demonstrated by solving problems in different application domains. The entirety of the identified practical problems forms the requirement base of SMILE; the solution of these problems will prove SMILE's feasibility.

The SMILE methodology can be seen from the meaning of its name: Semantic Metamodel-based Integrated Language Environment. The parts of this name have the following meaning in the scope of SMILE.

Semantic refers to the underlying idea to represent the semantics of languages and language constructs in an explicit way. This builds on the existing language description frameworks (e.g. grammars) and extends then to cover all aspects of modern computer languages.

Metamodel-based means that the approach is relying on modelling on all levels. We do not only use models for the actual systems, but also for the languages to be used for describing the systems and even for languages describing languages, etc. The

[1] The project is co-funded by the European Commission DG Information Society and Media, under the contract IST-004526.

important point is that the high-level descriptions of systems and languages should be complete. Our approach is *Integrated*, because we handle all the information in the same way, be it languages or systems. We describe the meaning of the constructs unambiguously in the (meta)model, and use this high-level description to generate tools.

Languages are in the focus of our project, because most of the activities can be related to languages. We are concerned with formal languages, and try to lift the level of formality so that all important aspects of the languages can be captured.

We build a complete *Environment*, such that the tools used within the SMILE project are finally built using the SMILE technology. A very important issue is that all parts of the SMILE description capabilities will be supported by tools: tools handling the description (reading, generating, editing) and tools transforming the description into other formats (e.g. code generation).

Currently we have succeeded in building the basic framework of SMILE according to [7] together with a simple user interface. Further work will include the creation of all the other components necessary for semantic modelling.

4 The Approach

The main idea is based on a combination of the accessibility rating according to EIAO with the modelling approach of SMILE. Modelling is especially important for web content, because accessibility requirements and web technology are constantly evolving. High level modelling of accessibility requirements can support more rapid generation of new test modules and improve the understanding of the accessibility barriers for web documents in a wireless environment.

UML models with OCL-like constraints [8] can be used to model requirements for accessibility testing [9]. Accessibility constraints will be attached to web document models that represent the relevant standards; such models guide the accessibility testing of web documents that are seen as instances of the mentioned models.

One approach to accessibility testing of XHTML documents was presented in [16]. The approach defines a three-level metamodel architecture where each level is to be represented in the SMILE-framework:

Metamodel. A simple metamodel that defines the most basic object-oriented concepts like class, property and composition (a small subset of the UML metamodel).

Model. A subset of the XHTML-specification [17] represented as a UML model instantiated from the metamodel. The model is extended with OCL constraints that define accessibility requirements to be fulfilled at the model instance level.

Model Instance. An XHTML-page is represented as an instance of the model. The OCL accessibility constraints are evaluated at this level to expose accessibility violations.

The EARL Evaluation and Report Language [10] developed by W3C, will be used for reporting deviations from standards and accessibility requirements.

The approach presented here is different from the one presented in [16]; it is meant to be more robust when it comes to deviations from the XHTML standard and it does not require a full model of the XHTML specification. The approach is based on the observation that an XHTML-document can be seen as a tree if we ignore external and internal links. Links could be handled with ordinary associations on the model level and links on the model instance level, but we have postponed this issue for a later version. By doing this simplification only the following object-oriented concepts are needed: class, property and composition.

The following subsections presents the constraint modelling technique by giving an example and how to implement the tool is also briefly described.

4.1 The XHTML-Document and the Accessibility Constraints

We used the three constraints defined in [16] as examples of accessibility constraints. These constraints are derived from the WCAG 1.0 Guidelines [11]:

1. Each image has to have a valid alt tag associated with it to provide text equivalents.
2. Only one of HTTP-EQUIV or NAME is allowed as attributes to a META tag.
3. The color of the text should have enough contrast with the background color. Checkpoint 2.2 in the WCAG says that it is important to ensure that the foreground and background color combinations provide sufficient contrast when viewed by someone having color deficits or when viewed on black and white screen.

These constraints are formulated in OCL as follows:

1. *Context* Img *Inv*: *libAcceptableAltTag*(alt)
2. *Context* meta *Inv*: name.size() > 0 xor http-equiv.size() > 0
3. *Context* font
   Inv: *libAcceptableContrast*(textColor, *getBody*(this)->any(true).background)
   -- We need an auxiliary recursive function that finds the body for this definition.
   Context block
   def: *getBody*(b : Block) : Set(Body) =
   if b.body->size() = 1 *then* body *else* b.*getBody*(composite) *endif*

The functions libAcceptableAltTag and libAcceptableContrast are library functions defined in the framework. We use the web page with the following XHTML code as example:

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN"
"http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">
<html xmlns="http://www.w3.org/1999/xhtml" xml:lang="en" lang="en" >
<head>
<meta http-equiv="Content-Type" content="text/html;charset=utf-8" />
<meta name="generator" content="gedit" />
<title>Test</title>
</head>
<body bgcolor="yellow">
<img src="http://www.eiao.net/kickoff_s.jpg" alt="" />
<br />
<font color="black">Participants at the EIAO Kickoff meeting in
Grimstad,
2004-10-14.</font>
</body>
</html>
```

The first constraint will expose that the document has an empty alt-tag which is in conflict with the WCAG 1.0 Guideline 1. The second constraint will test the instances of meta and find that this constraint is fulfilled. The third constraint refers to the font context; the call *libAcceptableContrast*("black","yellow") will give the result true; so this requirement is fulfilled for the font-element in the document.

4.2 The Three Level Metamodel Architecture

Inspecting an XHTML-document, or more generally an XML-document, reveals that you have elements, some elements have attributes and some have content (elements inside another element). In our approach an XML-document is related to a model instance and a model in the following way:

XML Element. An element in an XML-document is seen as an object of a class with the same name as the element-name.

XML Element Attribute. An element attribute in an XML-document is seen as a slot with the same value as the value specified for attribute; the slot is seen as an instance of a property with same name as the attribute; the property is a property of a class with the same name as the element containing the attribute.

XML Element Content. The content of an element is in a composition-relation to the containing element; the content is seen as part of the element.

There is no special translation of hyperlinks. Fig. 1 shows how the chosen XHTML–document is represented as a model instance, also the model and metamodel is shown.

The translation schema demands a metamodel that can support class, property and composition. The presented metamodel fulfils this demand and is an extremely simple reflexive metamodel[2].

Multiplicity is not defined; we simply see the multiplicity as fixed to zero-to-many on the part side of a composition and zero-to-one on the composite side of the composition (this can be specified as well-formedness rules in OCL). Constraints can be attached to classes in the form of text, e.g. constraint number two (described above) will be attached to class Meta on the model level! The metamodel defines that compositions are named; this is ignored at the model and model instance level because no names are given.

4.3 The Process of Defining and Evaluating the Constraints

The metamodel is defined in advance and it is this metamodel that dictates the translation schema and conceptually it defines our "world view". Obviously the metamodel does not have to be in place (built as a complete level in the framework) unless it is accessed which is not the case in the presented approach; as far as we know other metamodelling frameworks would require that the meta-model is in place, but not the SMILE framework.

On the model level only the classes, properties and compositions referenced by the constraints need to be modelled by the user, e.g. the user only needs to define class *meta* with attributes *name* and *http-equiv* if constraint two (see above) is the only one to be defined.

---

[2] A metamodel is reflexive if it defines all the entities it uses for its own definition.

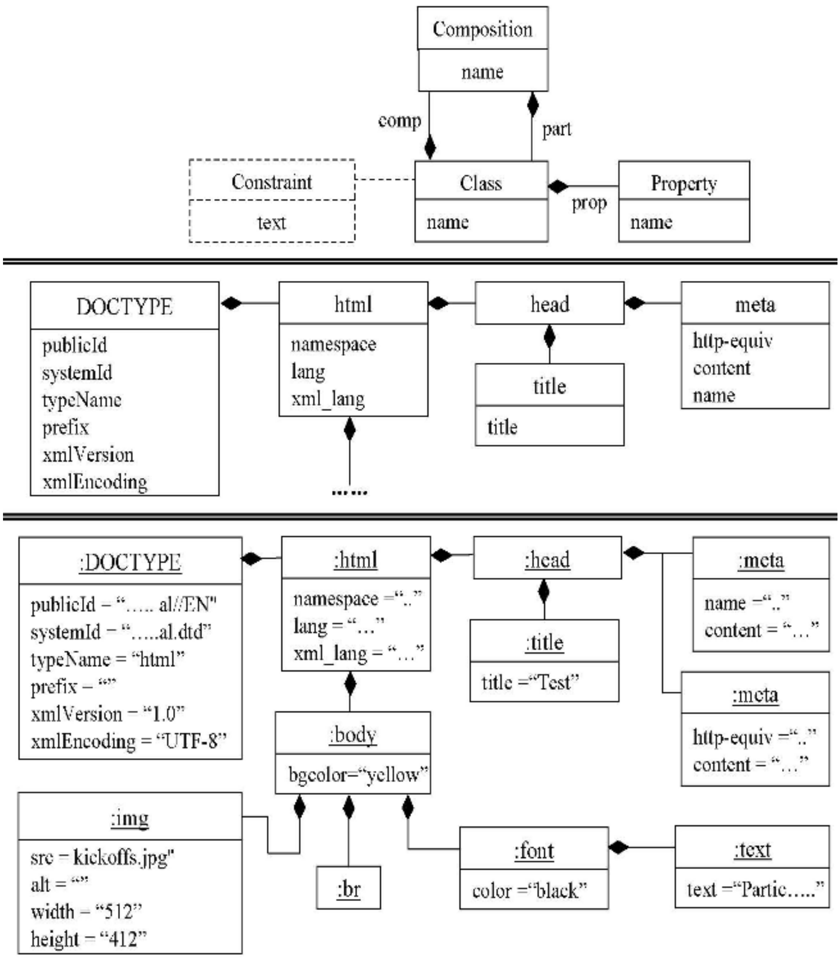


Fig. 1. The Metamodelling Architecture Presented with the help of UML (the model level is only partly shown)

We have developed a reader that reads an XHTML-document and converts it to a model instance as defined above; the reader is at the same time building a model, e.g. an element named *meta* gives a class *meta* on the model level as described in the subsection above.When the reader has done its job, the predefined model (which contains the constraints) is merged with the model that has been automatically created; at this point the constraints are tested on the model instance and a report describing the accessibility violations is created.

4.4 Benefits

Modern techniques for doing semantic validation of XML data are in nature declarative and typically based on XML-technology, e.g. Schematron [13] based on

XPath, and xLinkit [14] based on XPath. Formalisms based on XSLT and XPath are considered hard to write, understand and maintain (see [12]).

Schematron is more suitable for simple rules. Therefore we used Schematron in the first release of EIAO for expressing the basic building blocks of the accessibility tests. In order to express the higher-level tests, we will use a higher-level formalism, probably based on OCL.

[12] was published at the same workshop as [16], it describes a general approach for semantic validation of XML data based on metamodelling; the paper describes some strengths that also apply to our approach. Please find below a non-exhausting list of the advantages.

- Only one formalism is used (UML/OCL) and the constraints are specified at a high level of abstraction.
- It is partly graphical (UML).
- Deviations from the XHTML standard are allowed.

Some more benefits are added by our approach:

- We allow testing of XHTML-documents against constraints on fragments of models, i.e. we do not need the full specification of the XHTML standard.
- The parts of a document that conform to common deviations can be tested for accessibility; this is done by including model elements describing common deviations and attaching accessibility constraint to those elements.
- We apply special library functions in the OCL accessibility constraints, e.g. functions for testing foreground against background colors.

Our project is at the time of writing not yet completed (some important parts of OCL are not yet supported). In particular we want to investigate extensions of the navigation possibilities in OCL, e.g. "navigating through" objects of types that were unknown at model time (deviations from the standard), e.g. looking for background colors defined by elements of unknown type. However, we think we can already claim with some confidence that this approach to web accessibility assessment allows a flexible description of accessibility criteria and their evaluation.

5 Conclusion

Accessibility requirements and web technology are constantly evolving. High level modelling of accessibility requirements can support more rapid generation of new test modules and improve the understanding of the accessibility barriers for web documents. UML seems to be the natural choice for this modelling. The basic representation and instantiation technique from the SMILE project is very useful when it comes to the implementation of these ideas.

Starting from the prototype tool we have created, we will extend the subsets of XHTML and WCAG 1.0 covered. We will integrate a complete OCL interpreter into the SMILE framework, such that it is possible to express more constraints than just simple comparisons. Moreover, we will extend the library of functions needed to do sensible checks for accessibility.

This work was the result of successfully bringing together two relatively unrelated projects at AUC. The synergies between the SMILE and EIAO projects proved fruitful for the high-level description of accessibility requirements.

References

1. http://www.eiao.net
2. http://osys.grm.hia.no/smile/index.htm
3. http://osys.grm.hia.no/mebacc/
4. http://www.support-eam.org/
5. http://bentoweb.org/
6. http://www.w3.org/WAI/
7. J. P. Nytun, A. Prinz and A. Kunert: Representation of Levels and Instantiation in a Metamodelling Environment, NWUML 2004
8. Object Managment Group (OMG) Editor. OMG Unified Modelling Language. Specification, Version 2.0. OMG Document. OMG, http://www.omg.org, 2003
9. Tveit, Gjøsæter, Nytun: A UML Model of HTML for accessibility testing of web documents. Workshop on Web Accessibility and Metamodelling, 2005, Grimstad, Norway
10. http://www.w3.org/TR/EARL10/
11. http://www.w3.org/TR/WCAG10/
12. Dan Chiorean, Maria Bortes and Dyan Corutiu : Semantic Validation of XML Data A Metamodelling Approach, NWUML 2005
13. Rick Jelliffe: The Schematron Assertion Language 1.5. Specification, Academia Sinica Computing Centre, Oct. 2002; http://xml.ascc.net/resource/schematron/Schematron2000.html
14. C. Nentwich, L. Capra, W. Emmerich and A. Finkelstein: xLinkit: a Consistency Checking and Smart Link Generation Service, ACM Transactions on Internet Technology, 2(2), May 2002, http://xml.coverpages.org/xlinkitwp200102.pdf
15. http://www.wabcluster.org/uwem05/
16. Terje Gjøsæter, Jan P. Nytun, Andreas Prinz, Merete S. Tveit: Accessibility testing XHTML documents using UML, NWUML 2005
17. XHTML™ 2.0, W3C Working Draft 2005, http://www.w3.org/TR/xhtml2/
18. http://www.wabcluster.org

Scientific PDF Document Reader with Simple Interface for Visually Impaired People

Toshihiro Kanahori[1] and Masakazu Suzuki[2]

[1] Research and Support Center on Higher Education for the Hearing and Visually Impaired, Tsukuba University of Technology, 4-12-7 Kasuga, Tsukuba, Ibaraki 305-8521, Japan
kanahori@k.tsukuba-tech.ac.jp

[2] Faculty of Mathematics, Kyushu University, 6-10-1 Hakozaki, Higashi-ku, Fukuoka 812-8581, Japan
suzuki@math.kyushu-u.ac.jp

Abstract. We proposed our integrated system for scientific documents including mathematical formulae with speech output interface [1], named "ChattyInfty." This system consists of recognizer, editor and speech interface for scientific documents. Using this system, visually impaired people can read printed scientific documents with speech output. In this paper, we propose a new additional function of this system, which recognizes PDF documents. Using this function, visually impaired people can also read PDF documents including mathematical formulae with ChattyInfty. If PDF documents have embedded text, our PDF recognition engine does not only recognize PDF documents as page images but also utilizes the text information to verify its recognition results, to improve the recognition results. Recent commercial OCR software usually provides PDF output. Combining it with our PDF reader, users can get integrated recognition results of any such commercial OCR with Infty's results of mathematical formulae. This system also provides simple interface customized for visually impaired people. It enables them to read scientific document (as image files / PDF) in minimum steps of key operation.

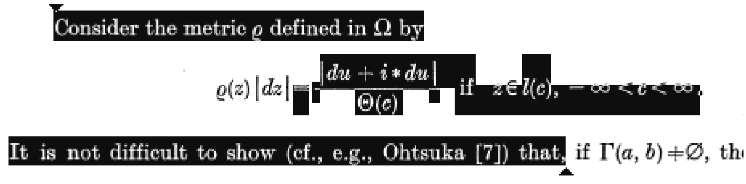
1 Introduction

A lot of scientific documents are provided as in PDF. If a PDF document has embedded text, visually impaired people can read it. However, some components of a document, for example, chemical formulae, mathematical formulae, and so on, often do not have proper embedded text information, which are very important parts of scientific documents. For example, mathematical formulae are recognized as ordinary text and they are given cryptic text (table 1), or they are either skipped and have no hidden text (table 2). In the both cases, the mathematical information can not be read out by screen reader.

We are developing a scientific document reader with speech interface, named "ChattyInfty" (fig. 1). ChattyInfty can recognize not only ordinary texts but also mathematical formulae. The recognition results can be output into LaTeX,

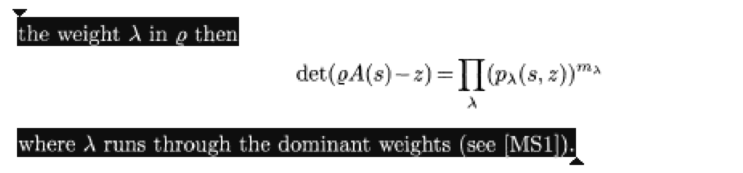
K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 48–52, 2006.

Table 1. The hidden text for mathematical formulae does not have correct mathematical information. The black painted parts have hidden text in the upper area. In the lower area, the hidden text is shown.




```
Consider the metric Q defined in 0 by
\du + i*du\ .. _. .
= J Q^T L if Z£|(c), -oo.<c<oo.
It is not difficult to show (cf., e.g., Ohtsuka [7]) that,
```

Table 2. The mathematical formula on the middle line is skipped and there is no hidden text for the formula




```
the weight A in g then
where A runs through the dominant weights (see [MSI]).
```

MathML, HTML, text embedded PDF (via LaTeX with dvipdfm) and so on, and can be edited with speech navigation which reads aloud not only ordinary text but also mathematical symbols and structure in principle and assign manners of aloud-reading of them. Using this system, visually impaired people can access to printed scientific documents and authorize them.

In this paper, we propose a new additional function of ChattyInfty which can recognize text embedded PDF documents and add mathematical formulae information to them. From a text embedded PDF document, its image file is generated and its text information is extracted as a text file. Then, the image file is recognized by ChattyInfty by utilizing the text information to correct the recognition results. We can also combine ChattyInfty with other OCR system which can output recognition results in text embedded PDF, and get higher quality recognition results. Using this system, we can add mathematical information to PDF documents and access them with speech navigation (fig. 3).

ChattyInfty provides simple interface to recognize and read PDF documents. Using the interface, a user can access to printed or PDF documents in the following steps:

1. selecting document files (PDF or page images) to recognize on Windows Explorer,
2. opening Context-Menu and selecting 'Recognize by InftySystem' (fig. 2).

Then the selected documents are recognized, and the recognition results are opened by ChattyInfty. A setting dialog application is also provided to change some parameters (language and resolution of images) for the recognition (fig. 2). These interface and application are made of standard Windows components, so general screen reader can navigate them.

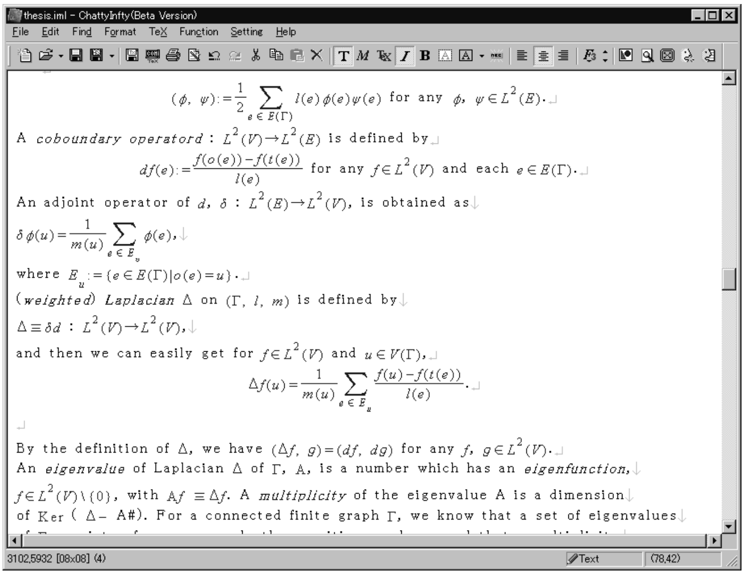


Fig. 1. Snapshot of ChattyInfty

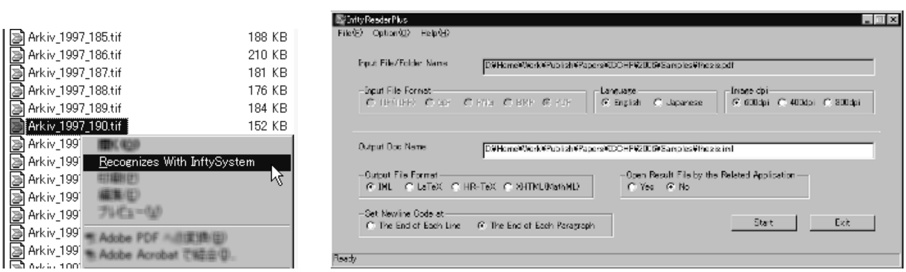


Fig. 2. Context-Menu for recognizing by Infty (left) and the recognition interface for advanced recognition (right)

2 Outline of PDF Recognition

This system recognizes a text embedded PDF document as follows (fig. 3);

1. extracting a text information from a PDF document as a text file,
2. converting the PDF document to a Multi-TIFF image file,
3. recognizing the image file by ChattyInfty,
4. correcting the recognition results by matching with the extracted text information,
5. outputting the recognition result file in our original XML format and converting it into other formats selected by user.

The detail of this method is described in [2].

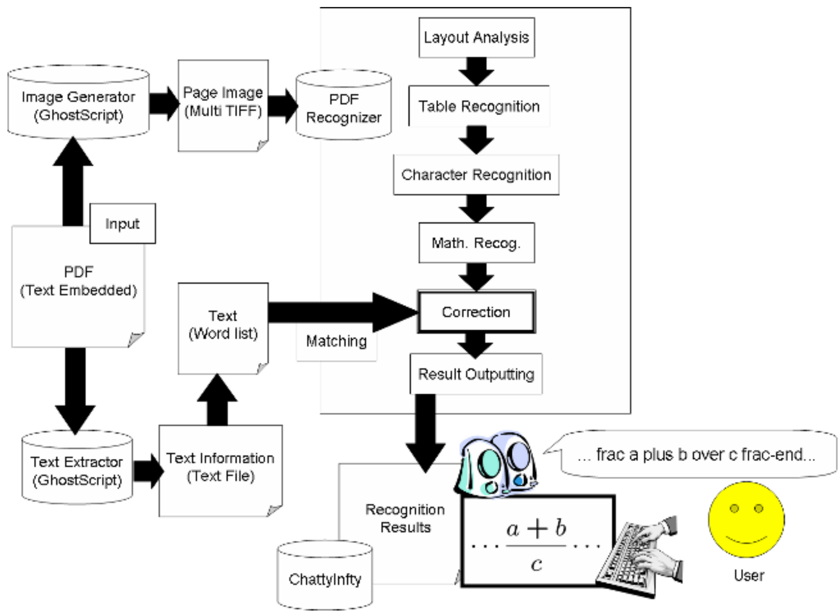


Fig. 3. Recognition process of PDF documents by ChattyInfty

3 Future Plans

In this paper, all information in PDF – font information, special characters, bounding rectangle, etc. – could not be used, because only embedded text are used. More improvement can be expected by utilizing document information directly from PDF files.

This system uses other OCR system via PDF files. It is probably possible to integrate the best recognition results of several OCR systems, or to update digital libraries whenever OCR systems are upgraded with the advance of technology.

References

1. M. Suzuki, T. Kanahori, N. Ohtake and K. Yamaguchi, "An Integrated OCR Software for mathematical Documents and Its Output with Accessibility", *Computers Helping people with Special Needs*, 9th International Conference ICCHP2004, Lecture Notes in Computer Sciences 3119, Springer (2004) pp.648–655
2. T. Kanahori and M. Suzuki, "Refinement of digitized documents through recognition of mathematical formulae", to Appear in the Proceedings of the 2nd International Workshop on Document Image Analysis for Libraries, 6pages.

English “Oblique” Listening System – Rapid Listening System for the Blind and Visually Impaired, and Its Evaluation

Shinichi Torihara[1], Miyoko Nakamura [2], Nami Ueda [2], Takuya Wada[3], and Shun Ishizaki[4]

[1] Graduate School of Media and Governance, Keio University, 5322 Endo Fujisawa, 252-0816 Japan
torihara@sfc.keio.ac.jp
[2] Keio Research Institute at SFC, 5322 Endo Fujisawa, 252-0816 Japan
[3] Graduate School of Media and Governance, Keio University, 5322 Endo Fujisawa, 252-0816 Japan
[4] Faculty of Environmental Information Graduate School of Media and Governance, Keio University, 5322 Endo Fujisawa, 252-0816 Japan

Abstract. We propose an "oblique" listening system in the English language for the blind and visually impaired by controlling the speed of Text-to-Speech based on parts of speech in which the important parts are synthesized 'relatively' slower while the unimportant at the maximum speed. In the system evaluation experiment, English natives with vision impairments were required to listen to three short passages of Text-to-Speech at the three types of speed and answer the questions of comprehension after the base-line speed was calculated from the measured maximum speed of recognizing 'a word' and that of recognizing 'a sentence' for each subject. We controlled the Text-to-Speech speed: base-line, 'oblique listening' speed, and simple high speed, by which the duration is equal to the one by our system but without any speed variation. The results show that the “oblique" listening system is better than a simple high speed system even though not exceeding the base-line.

1 Introduction

Normally sighted readers skip more words if they are speed-reading or skimming. The findings from the eye tracker study show that readers make longer pauses at points where processing loads are greater, which implies the important parts of the text. According to the reading literature, almost all content words are fixated and the short function words tend not to be fixated [1]. However, the blind and visually impaired must rely on a conversion from text to speech. The problem with this is that speech sounds generated from text are sequential and linear. If one listens to speech at speeds faster than the average spoken rate, one will not be able to understand it very well [2]. We propose an "oblique" listening system in the English language for the blind and visually impaired. In general terms, our proposed method works by controlling the speed of Text-to-Speech based on the different parts of the text. The important parts of the text are synthesized to speech at slower rates, while the less important

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 53 – 60, 2006.

parts of the text are sped up to the maximum speed of the system. In this paper, we will describe English "oblique" listening system, and report the results of its system evaluation taken by English natives with vision impairments.

2 "Oblique" Listening System

2.1 Purpose of the System

In obtaining the information by listening, it is very difficult to listen to and understand much information at the same time because speech sounds are sequential and linear. Then we have to speed up the rate of the text by TTS(Text□to□Speech) to obtain much more information. The faster the text needs to be played, the more difficult it will be for one to understand it [3]. The persons with vision impairments need such an effective spoken language information processing system by auditory interface as the "oblique" listening system.

2.2 Characteristics of the English Language

In English, sentence stresses are located in a sentence. It takes equal time to pronounce from a stressed word to its following stressed word. This is what we call "Stress-to-Stress Isochronism" [4]. The words with the stress are spoken stronger and slow, and are important in most cases. We apply for the characteristics to control the speeds.

2.3 Specification of Important Words

We define important words as follows [5];

(1) Noun
(2) Verb
(3) Predicative adjective
Your daughter is *pretty*.
(4) Expression of number and quantity
He took *a few* biscuits.
(5) Negative expression
The girl *isn't* a student.
Nobody contributed more to the understanding of dreams than Freud.

2.4 Configuration of the System

The system for "oblique" listening consists of following two programs. The input parameters of the program are (a) text file specified with '[', ']' for important words (b) maximum speed of recognizable words for each subject(unit is WPM(word/minute)) (c) maximum speed of recognizable sentences for each subject(WPM). The output file includes words tagged with calculated speed rate. The input of the TTS program is text file embedded in speed tags. The program synthesizes through MSAPI (Microsoft Speech API) and stores the result as wave file. We use English TTS engine, IBM OutLoud for our system. We will show the system configuration in Fig. 1.

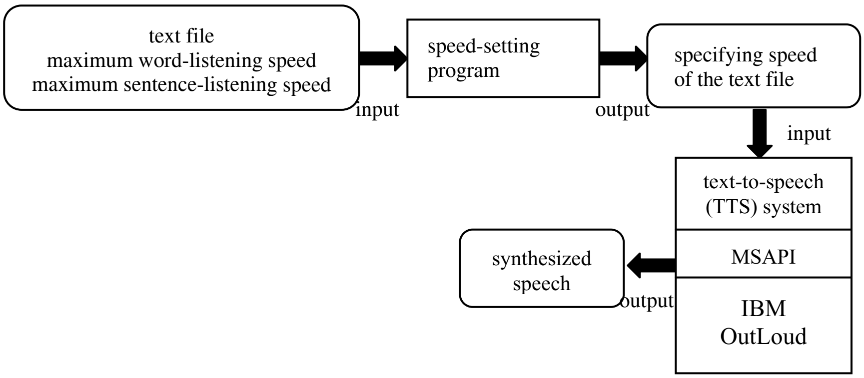


Fig. 1. System configuration

3 Measurement of the Recognizable Speed

Great individual differences are observed for listening to the human speech or Text-to-Speech (TTS) at high speed. Especially, the visually impaired are reported to have a big difference of recognizable speed according to their experience of being habituated to listening to TTS or not [6]. In our study, we decided to measure two kinds of the recognizable speed for words and sentences because the system is aimed at listening comprehension of both sentences and important words in the text.

3.1 Subjects

Subjects were 20 English native speakers: 12 visually impaired who used TTS system almost everyday, that is, who had experience of using the system of two or more years for one or more hours per day and 2 visually impaired who did not use such a system and 5 normally sighted. We call the former "experienced users" and the latter "standard users." The ages of subjects were as follows: 3 twenties, 4 thirties, 10 forties, 2 fifties, 1 sixties.

3.2 Measurement of the Maximum Speed of Recognizable Words

Selection of the Words used for the TTS System. 200 frequent words were selected from "Core Vocabulary of International English" [7]. Specifically, among the content words, one-syllable words, two-syllable ones, and three-syllable ones were selected from those with the highest frequency to sum into 200. Considering the balance of each part-of-speech used in our system, 42 adjectives, 92 nouns, and 66 verbs were selected, which consist of 60 one-syllable words in total (7 adjectives, 31 nouns, and 22 verbs), 79 two-syllable words in total (14 adjectives, 34 nouns, 31 verbs), 61 three-syllable-and-more words in total (21 adjectives□27 nouns, 13 verbs).

Procedure. The TTS materials were used in the measurement in the manner of 20 gradually lowering speed stages from 800WPM (800 words are spoken per minute) to

420WPM with the interval of 20WPM. We used personal computer, IBM ThinkPad Windows2000, and the subjects were instructed to hear the TTS through the headphones.

In order to measure the approximate maximum as the 1st phase, we used one word with two syllables at each speed one by one from the highest speed. Then the subjects were asked to speak out the spelling of the word that was heard, and the correction was judged. When the subjects answered correctly, the subject heard 4 English words at the 20WPM higher speed and the correction was judged again, taking the error for the test of the 1st phase with only one word at a time into consideration. Since the 4 words have variation with one to five syllables, 4 English words at 20WPM higher speed were used. When the subject answered to the 4 words correctly, the same procedure was repeated at the 20WPM higher speed. Thus we define the maximum speed that the subject can speak out the spelling of the word correctly as the Maximum Speeds of Recognizable Words.

Result of the Experiment. As for the average of the maximum speeds of recognizable Words, the significant difference (T test p> 0.0255) between the visually impaired group (701WPM) and the normally sighted group (592WPM) is seen together with the significant difference (T test p> 0.0009) between the "experienced users" (722WPM) and the "standard users." (586WPM)

3.3 Measurement of the Maximum Speed of Recognizable Sentences

Material. 100 frequent words, which are defined as concreteness nouns, were selected from "Core Vocabulary of International English"[7] regardless of the number of syllables from the highest frequency. We asked the subjects to listen to English sentences one by one and tell us what the sentence describes by choosing one word from the 4 choices. 3 choices except the correct answer are selected from the reference [8]. Each choice has some similarity with the correct answer.

Procedure. The TTS materials were used in the measurement in the manner of 20 gradually lowering speed stages from 600WPM to 220WPM with the interval of 20WPM. We asked the subjects to choose one word from the 4 choices. The reason we set 600WPM at the highest speed is because the speed is regarded as somewhere near the maximum speed of recognizable sentences by preliminary experiment.

When the subjects answered correctly, as we did in the measurement of the maximum speeds of recognizable words, we presented the sentence at the 20WPM higher speed to the subjects and asked them to speak out the correct word which describes the sentence choosing from the four answer choices given by the experimenter. Since it was well predicted that the sentence yields a little difference due to the easiness of comprehension, the subject heard 4 English sentences at the +20WPM higher speed and the correction was judged again. When the subject answered to the 4 sentences correctly, the same procedure was repeated at the 20WPM higher speed. Otherwise the same procedure was repeated at the -20WPM lower speed. Thus we define the maximum speed that the subject can choose the right choice of the sentence correctly as maximum speed of recognizable sentences.

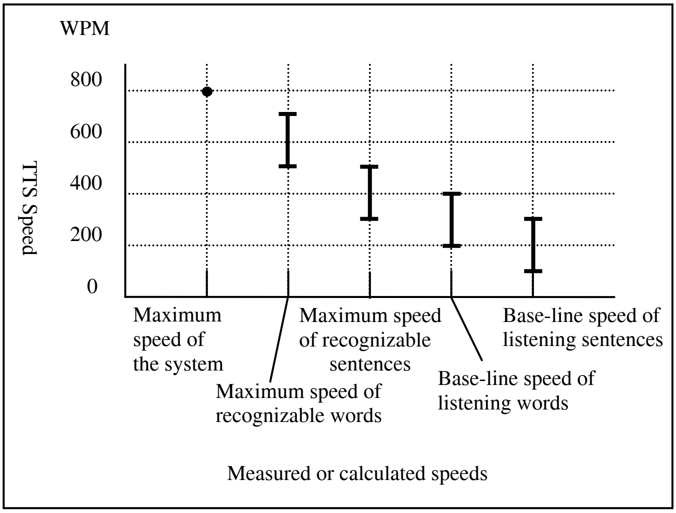


Fig. 2. Distribution of the measured maximum speeds and the base-line speed

Results. As for the average of maximum speed of recognizable sentences, the significant difference (T test p> 0.0007) between the visually impaired group (569WPM) and the normally sighted group (472WPM) is seen together with the significant difference (T test p> 0.0001) between the "experienced users" (578WPM) and the "standard users." (483WPM)

3.4 Definition of Speed for Three Systems

We define that base-line speed is easy, comfortable and efficient speed for each subject. Base-line speed is a standard or measurement and is used for comparison in the other speeded-up systems. The slowest speed in maximum speed of recognizable sentences the "standard users" was 440WPM. The speed, 440WPM divided by two is similar speed to the default speed (180WPM) of the TTS engine. Therefore, we define that base-line speed is maximum speed of recognizable sentences for each subject divided by two [9]. That is, base-line speed differs from subject to subject. In other words, the twice speed where each subject is usually listening to text is maximum speed of recognizable sentences for each subject. For "oblique" listening system, we use two kinds of speed to synthesize. The faster speed is maximum speed of the TTS engine (800WPM). The slower speed of important portions is defined as maximum speed of recognizable words for each subject divided by two. This definition is the same reason as we define base-line speed. The simple high speed system has the same duration or playback time as the corresponding "oblique" listening system.

4 Evaluation Experiment of the Oblique Listening System

Materials. The speech to which subjects listened consists of three short TTS passages, 400 words, and the 6 recorded questions 4 of which include 3 answer choices spoken by an English-Japanese bilingual. All the passages had such characteristics as requiring no expert knowledge but allowing no answer inferred by common sense, and selected mostly from the textbooks of English as the second language. The number of passage presented for each subject was decided to be three in consideration of the cognitive load and fatigue of a subject. The following three text types, 1) autobiography, 2) expository, and 3) narrative, were selected in order to avoid the inclination suggested by the findings from the pretest. 6 questions for every three passages were asked to the subjects in order to examine their comprehension after presenting every TTS passage, The first and the second questions were TRUE/FALSE questions which examined whether the sentence fit for the content of the passage. Each of the third and the fourth question was consisted of one sentence extracted from the passage and two more sentences in which one target word was replaced by another different word in the different category, that is, 3 answer choices in total, and subjects were asked to choose one of them. In the fifth and the sixth questions, there were one sentence which was not completely the same as the one in the passage but had almost the same meaning and two other sentences the meaning of which was different from the content of the passage, and subjects were asked to choose the correct sentence which fit the content of the passage from those three choices. All the third, fourth, fifth, and sixth questions had no target word which was realized rapidly by the oblique listening speed control.

Procedure. Experiment was individual testing carried out in the quiet room with the presence of experimenters. All the questions and choices for the text comprehension test and the instructions, which were spoken by an English-Japanese bilingual, were presented auditorily from the headset attached to the personal computer as well as the TTS passages realized by our system. Subjects were asked to listen to every single passage just once and twice to any single question and choice. The presentation order of three passages, questions, and choices was the same among all the subjects, but three types of speed, such oblique listening, as base-line, and simple high speed, the last two as control conditions, and three kinds of passage were counterbalanced among the subjects and the control was intra-subject which required them to listen to any of the three TTS passages at any of the three types of speed.

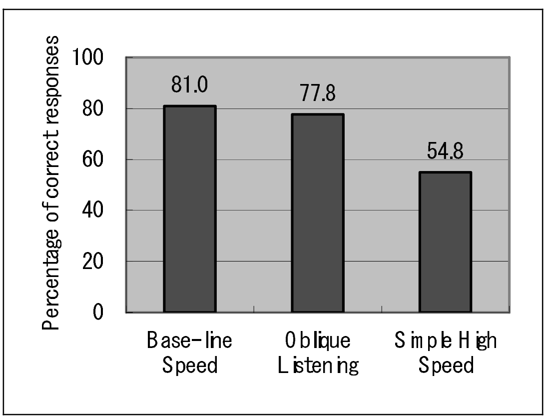


Fig. 3. Percentage of correct responses to Passage 1

Results. Concerning each passage, the comparison among the intra-subject conditions, by comparing the mean percentages of correct responses to each type of speed, shows a significant difference between "oblique listening" condition and that of "simple high speed" condition in Passage 1(one-tailed t test, p=0.0491). (See Fig. 3)

When comparing the mean percentage of correct responses to each passage, there is a significant difference between Passage 1 and Passage 3 (one-tailed t test, p=0.0147)and another between Passage 2 and Passage 3 (one-tailed t test, p=0.0185)

To all the passages, there is no significant difference both between the percentages of correct responses of the visually impaired group (15 subjects) and those of the normal sighted group (5 subjects) and between that of the experienced users and the standard users. Also, concerning every passage, the subjects with good performance and those with poor performance tend not to be in the same group by chance.

5 Discussion

In Passage 1, we can find the significant difference between "oblique" listening system and simple high speed system. In Passage 3, in the other hand, although we cannot find the significant difference between the two, the mean percentage of correct responses of simple high speed system is slightly higher than that of "oblique" listening system. The reason is why Passage 3 must be relatively easy to understand. From the characteristics of passage, Passage 3 is a narrative text or story. It is assumed that it would be easy for the subjects to understand and memorize the semantic content out of the passage. Throughout the passages, we cannot find the significant difference between the subjects with vision impairments (15 persons) and the sighted subjects (5 persons). The reason is that the systems adjusted to each subject in speed setting. We consider that the difference of speed doesn't lead to that of understanding. Therefore, although the difference of speed is significant, the difference of the results does not occur after the normalization. We consider that it is very adequate to assign speed to TTS engine after measuring maximum speed of recognizable sentences and words for each subject.

6 Conclusion

We proposed an English rapid listening system for the blind and visually impaired by control-ling the rate of TTS based upon the different parts of speech. They feel the need for an improved system in order for them to gain access to widespread interdisciplinary knowledge. Our proposed system will help them benefit from the advantages of so-called "diagonal" reading currently available only to sighted people able to read print.

References

1. Just, M.A., Carpenter, P.A.: A Theory of Reading. From Eye Fixations to Comprehension. Psychological Review Vol. 87 (1980) 329-354
2. Asakawa, C., Takagi, H., Ino S., Ifukube, T.: The Highest and the Most Suitable Listening Rate for the Blind in the Screen Reading Process, Proc. Human Interface Symposium 2002 (2002) 543-546

3. Vemuri, S., DeCamp, P., Bender, W., Schmandt, C.: Improving Speech Playback Using Time-Compression and Speech Recognition. Conference on Human Factors in Computing Systems, Proceedings of the SIGCHI conference on Human factors in computing systems, Vienna, Austria (2004) 295-302
4. Brazil, D.C., Coulthard, C, Johns.: Discourse Intonation and Language Teaching, Longman, London (1980)
5. Quirk, R., Greenbaum, S., Leech, G., Svartvik, J.: A Comprehensive Grammar of the English Language. Longman Group UK Ltd., Harlow (1985)
6. Asakawa, C., Takagi, H., Ino S., Ifukube, T.: The Optimal and Maximum Listening Rates in Presenting Speech Information to the Blind. J. Human Interface Society, Vol. 7 No.1 (2005) 105-111
7. Peyawary, A. S.: The Core Vocabulary of International English; A Corpus Approach. The SIT Centre, Bergen (1999) 150-176
8. Longman Dictionary of Contemporary English, Third Edition. Kirihara Shoten (2002)
9. Watanabe, T., Sakajiri, M., Sashida, C., Okada S. : A Survey of Windows Accessibility by Visually-Impaired PC Users. Technical Report of the Institute of Electronics, Information and Communication Engineers (IEICE). Vol. 100 No.256 (2000) 37-42

Creating Pictures by Dialogue

Ivan Kopeček and Radek Ošlejšek

Faculty of Informatics, Masaryk University
Botanicka 68a, 602 00 Brno
Czech Republic
{kopecek, oslejsek}@fi.muni.cz

Abstract. This paper deals with the problem of generating a picture of a scene by means of describing it in the terms of natural language ontologies. The corresponding graphical format, being formally represented in the form of the Pawlak information system, involves a full semantic description of the picture. Therefore, the pictures generated in this way "blind-friendly", i.e. they can be automatically fully described. A simple example of such a picture generation is presented here.

1 Introduction

Developing methods and algorithms that could help visually impaired people to manage, recognize and process graphical data is a challenging problem and simultaneously an important task. Up to now, most of the approaches (e.g. [2,3]) to this problem deal with tactile graphics. However, blind users have to cope with electronic graphics on the web. Web information is one of the most important information sources and graphical data constitutes a substantial part. To create and recognize pictures may seem to be too difficult a task for blind users, but in practice we can often see that the will, determination and ability of blind users overcome many complicated obstacles. Nevertheless, graphics still presents a major problem.

In this paper, we describe and discuss the idea of developing graphics through dialogue, in a way that is fully accessible to blind users. Simultaneously, the graphics developed in this way is encoded by a format which enables the picture to be described. Therefore, the format of the picture provides simultaneously the information describing the picture.

Why should a blind person try to create pictures? One reason may follow from the need to illustrate personal web pages. Another case is when the blind users want to send an illustrated message or e-mail, e.g. birthday congratulation, to their seeing friends. Because our pictures are self-describing by their format, they are convenient for all blind-friendly web pages and they are also suitable for blind-friendly computer games. And the authors are sure that the visually impaired will find much more other applications.

2 Modeling Picture Description by Means of Pawlak Information Systems

We apply the notion of information systems to describe the properties of the basic elements that we use to model dialogue communication [4]. Formally, information system (in the sense of Pawlak [5,7]) is defined as follows.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 61–68, 2006.

Let U, T, V be nonempty sets and let f be a mapping assigning to any pair of elements $u \in U$ and $t \in T$ an element from V. Then the ordered quadruple $S = (U, T, V, f)$ is said to be an information system. The elements of U are called objects, the elements of T attributes, and the elements of V values of attributes. For our purposes, we will assume that the objects will be uniquely identified by their names. Two objects u, v are called interchangeable if $f(u, t) = f(v, t)$ for all $t \in T$. It is assumed that the sets U, T, V are finite.

Using this formal model offers some advantages. First, we can apply the theory behind this general schema to solve relevant problems (e.g. to find the minimal set of attributes that distinguishes a set of objects [5]). Second, we have a simple and flexible scheme for further analysis. Third, the data in the form of Pawlak information system can be easily derived from most of the forms of ontologies. And finally, this formalism enables us to introduce a metrics to the set of objects which consequently gives us the possibility of utilizing sophisticated search methods in dialogue strategies.

Our main interest is in how to perform a dialogue for finding a suitable object. There are two basic strategies appropriate to this task.

(a) Users are asked to specify all values of all attributes. Then they can eventually give a new name to the object enlarging the set of all objects.
(b) Users are asked to specify the object by its name. This can be done either by the user initiative, directly specifying the object, or by the system initiative, reading some (or all) possible objects and asking the user to choose one.

It is clear, that both strategies can be used depending on what the user knows about the database of objects. If users are not sure whether the searching object is in the database or are not sure about its name, they will prefer the first strategy. If users know the name of the object, they will use the second strategy.

In many situations, both strategies can be efficiently combined. The idea is in "approximating" the searched objects by the second strategy (b) and then applying the first strategy (a).

To enable an efficient implementation of the mixed searching strategy, we introduce a metrics on the space U:

Let $u, v \in U$ and let $c(u, v)$ be the number of the attributes in which both objects differ (or, more precisely, in which the corresponding values of attributes differ). Then we define $d(u, v) = \sqrt{c(u, v)}$.

It is easy to see that the function d is symmetric, i.e. $d(u, v) = d(v, u)$ for every u, v and that it satisfies the triangle inequality, i.e. $d(u, v) \leq d(v, w) + d(w, u)$ for every u, v, w. Further, it holds $d(u, u) = 0$.

Generally, it need not be always true that $d(u, v) = 0$ implies $u = v$. This condition is violated if there are two different objects that are interchangeable. Therefore, if there are no different interchangeable objects, d is metrics, otherwise it is pseudometrics. In either case, the function d fulfils the substantial properties to be utilized in searching methods developed for searching in metric spaces [11].

Generally, the methods for searching in metric spaces are based on indexing the data by finding representants of data clusters. Based on the data structure, we can choose a suitable method. In our mixed dialogue strategy, representants of clusters are used to approximate the searched object.

Another feature enhancing the efficiency of the dialogue strategies is a suitable use of subobjects. Subobjects are objects that contain a special attribute value $not\_defined$. For instance, if the ontology describes a human being, the subobject $HEAD$ may be the objects for which all attributes not describing the head have the value $not\_defined$. Subobjects are used to enable a more structured approach in defining the object. The final generated object must not be a subobject.

Finally, to each object and each attribute, the set of consistent attributes and attribute values can be prescribed. For instance, for the objects $WOMAN$, the attribute beard is not consistent. This approach enables omitting the inconsistent values in the lists of attributes and attribute values for a concrete object.

Instead of going into more technical details, we present a very simple example: The task is to define and generate pictures of a fairy-tale characters known as $MANKA$ and $RUMCAJS$ (both these characters are well known by kids in Czechia; $RUMCAJS$ is a merry robber, something like Robin Hood, and $MANKA$ is his brave wife, see Fig. 1).

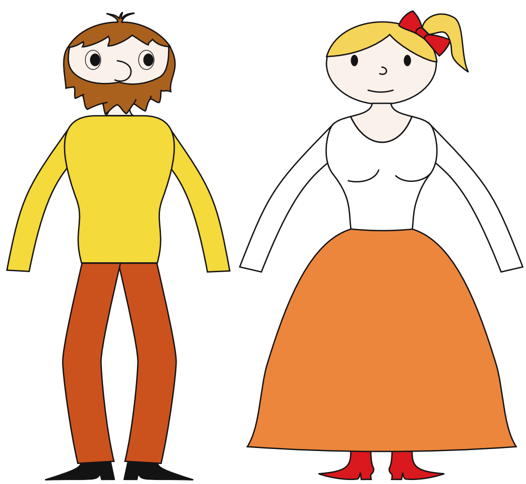


Fig. 1. Rumcajs and Manka

Assume a suitable ontology behind the corresponding system and let the parameters of the system be as follows.

Attributes: $body\_width$, $body\_height$, $breasts$, $nose$, $eyes$, $hair\_style$, $beard$, $trousers\_skirt$;

If $u \in U$, let $Val(u)$ be the set of the values of the attributes of u that differs from the value $not\_defined$.

$$Val(body\_width) = \{normal, thick, thin\}$$
$$Val(body\_height) = \{normal, tall, small\}$$
$$Val(breasts) = \{normal, small, big\}$$
$$Val(nose) = \{normal, small, big\}$$

$$Val(hair\_style) = \{baldheaded, rumcajs\_style, long\_hair, braid\}$$
$$Val(beard) = \{beardless, moustache, whiskers, fullbeard\}$$
$$Val(eyes) = \{normal, small, big\}$$
$$Val(trousers\_skirt) = \{trousers, skirt\}$$

Hence, for our example we get the following set of the attribute values:

$$V = \{\, normal, thick, thin, tall, small, big, baldheaded, rumcajs\_style, braid, beardless, moustache, whiskers, fullbeard, trousers, skirt, not\_defined\}$$

The attributes and their values define the objects of the system. We will assume, that the objects $WOMAN$, $MANKA$ and $RUMCAJS$ are among the objects of the system.

In this example we assume that the user is familiar with the basic notions (object, subobjects, representative, attribute etc.).

Dialogue fragment $RUMCAJS$ [1]:

```
S: Will you choose the name of the object or its attributes?
U: Name.
S: What is the name of the object?
U: RUMCAJS.
S: (System generates the picture)
```

Dialogue fragment $MANKA\_1$:

```
S: Will you choose the name of the object or its attributes?
U: Attributes.
S: Will you define subobjects, representative or all attributes?
U: All attributes.
S: Choose the body width. It can be normal, thick or thin.
U: Normal.
S: Choose ...
   (systematically asks all the attribute values and then
    generates the picture)
```

Dialogue fragment $MANKA\_2$:

```
S: Will you choose the name of the object or its attributes?
U: Attributes.
S: Will you define subobjects, representative or all attributes?
U: Representative.
S: Will you choose the representative by name or from the list?
U. By name.
S: What is the name of the representative?
U: WOMAN.
S: Will you make some changes?
U: Yes.
S: Will you define subobjects, representative or all attributes?
U: Subobjects.
```

[1] S = system, U = user.

```
S: Will you choose the subobjects by name or from the list?
U: From the list.
S: The subobjects are: Head, Body, and Dress.
   Choose one, please.
U: Head.
S: The attributes of the Head are: eyes, nose, hair_style.
   Choose one, please.
U: hair_style:
S: Standard hair_style for WOMAN is long_hair.
   You can change it to braid.
U: Change to braid.
S: Another changes?
U: No.
S: (generates the picture)
```

In this dialogue fragment, the representative $WOMEN$ was used to create $MANKA$. In our example, $MANKA$ differs from $WOMAN$ just in having braid; this is, of course, just to keep the example as simple as possible.

Notice that the system does not ask for the inconsistent attribute values (the attribute $beard$ and the attribute value $baldheaded$). This does not mean that it is not possible to use them for some objects, it just means they are not included in the standard lists.

3 Graphics Format

The basic concept for creating virtual scenes, including 2D pictures, is *composition*. Complex object is composed of more simple components which create a hierarchical representation, a so-called *scene graph* [6,8]. The parent–child relationship in the hierarchy does not mean only the relationship between a complex object and its logical component. The relationship also describes a geometric dependency: a child is transformed in the space relative to the parent. This geometric relations predestine an exact structure of the hierarchy which can differ from the hierarchical structure used by the dialogue system. For example, a human is composed of a body, head, legs etc. But the head is positioned relatively to the body and then it should be the child of the body object in a scene graph. We therefore need some generic structure that will enable us to build various hierarchies with minimal restrictions.

Virtual objects contain various characteristics. Apart from relative transformation, it is typically a shape (i.e. geometry) and appearance (e.g. color, texture, transparency, etc.). For the interconnection with the dialogue subsystem a virtual object should also have some kind of lexical description.

Fig. 2 shows a possible object-oriented decomposition. This diagram is a standard UML class diagram [1,9] analysing the static dependencies between classes. Generic class `GraphicsObject` covers any virtual object on a painted picture. `Composite` class represents the ramification in the scene graph (an object composed of more simple objects), `Primitive` class represents leaf objects with no children. Attributes describe detailed visual characteristics and can be associated with either complex object or its components. In the figure there are several examples of attributes with clear meaning. Note that these attributes are not directly related to the attributes of the dialogue system

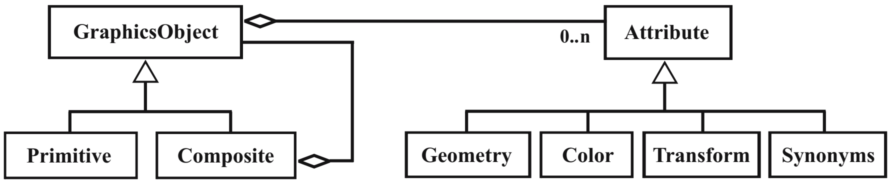


Fig. 2. Scene Graph Decomposition

discussed in the previous section. Graphics attributes correspond to either subobjects in the dialogue system or their attributes.

The configuration of classes in Fig. 2 corresponds to the COMPOSITE pattern [10] with all its good features that make the decomposition very simple but still general. We have the freedom to build complex objects bottom up. We could make a new complex object by selecting existing virtual objects, changing their attributes and collating them together in a proper hierarchy. A newly created object can be reused in the same way for building even more complicated objects.

4 Strategies for Building the Picture

A graphics subsystem has to be able to depict a picture as requested by a dialogue system. The picture requirements are specified by a set of lexical attributes that determine the appearance or even the presence of graphic components, e.g. the presence and color of a skirt in the case of female being. Every lexical attribute and subobject therefore has to have its graphical representation stored somewhere in a database. The task of the graphics subsystem is to select the desired objects from the database and possibly change their graphic attributes.

The second task of the graphics system is to collect the elements in a proper way. The concrete set of attributes taken from the dialogue system is related to a concrete type of picture. For example, the set of attributes that describes a human being is different from the set that describes a car. A graphics system therefore has to know how exactly build the final picture in each case, i.e. how convert the set of lexical attributes to the hierarchical scene graph.

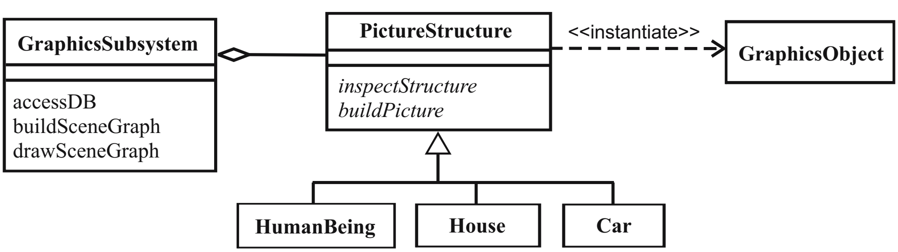


Fig. 3. Graphics Subsystem

Fig. 3 shows object-oriented decomposition of graphics subsystem using STRATEGY design pattern [10]. A `GraphicsSubsystem` class covers the main functionality of the graphics system, i.e. access to a database and building as well as visualizing a scene graph. Knowledge about the internal structure of the type of picture is moved to a separate class hierarchy represented by the abstract class `PictureStructure`. Concrete subclass provides information about the logical structure of the image type, the *inspectStructure()* method, as used by the dialogue system. In the figure there is a few concrete examples including a human being, house and car images.

The `PictureStructure` class is also able to build scene graphs on the basis of the requirement of the dialogue system. During the painting process the graphics subsystem forwards requests from the dialogue system to a concrete strategy. The strategy then interacts with the graphics subsystem in order to retrieve graphics components from database and builds the scene graph. Finally, the scene graph can be easily rendered on the screen.

5 Conclusions and Future Work

This paper has presented a brief outline of the underlying ideas and architecture of a system for generating pictures via a human–computer dialogue. This approach enables visually impaired people to create their own pictures that can be used in web presentations, emails, publications etc. Another benefit of such dialogue generation of pictures is that the corresponding graphical format contains a full description of the picture.

The generation of the graphics by means of dialogue has been discussed and tested by students and blind users during the 2005 course on Assistive technologies at Faculty of Informatics, Masaryk University Brno, and in cooperation with Teiresias Centre (Support Centre for Students with Special Needs) of Masaryk University. The results of the testing shows that the approach is promising and interesting.

Future work will be directed towards developing methods for enlarging and integrating basic ontologies, enhancing dialogue strategies and to implementing graphical systems for basic graphical primitives acquisition. An interesting problem we face is how to inform the user about a picture optimally. Another major task is the integration of modules that would add aesthetic principles to generated pictures.

Acknowledgement

The authors are grateful to James Thomas for reading a draft of the paper and to the students and staff of the Support Centre for Students with Special Needs of Masaryk University for their advices, support and collaboration. This work has been supported by Czech Science Foundation, Czech Republic, Contract No. GA 201/06/P247, by Ministry of Education of CR, Contract No. LC06008 (Center of Computer Graphics) and by Faculty of Informatics, Masaryk University.

References

1. Arlow J., Neustadt I.: UML 2 and the Unified Process, Second Edition. Addison-Wesley, 2005.
2. Intact, `http://www.surrey.ac.uk/~pss1su/intact/`.
3. Kaczmarek, K.: Electrotactile Display for Computer Graphics to Blind. Research Report 5-R01-EY10019-08, University of Wisconsin, 2004.
4. Kopeček, I.: Modeling the Information Retrieval Dialogue Systems. In *Proceedings of the Workshop on Text, Speech and Dialogue – TSD'99*, Lectures Notes in Artificial Intelligence 1692, Springer-Verlag, 1999, pp. 302–307.
5. Novotný, M., Pawlak, Z.: Algebraic Theory of Independence of Information Systems. Fundamenta Informaticae XIV, 1991, pp. 454–476.
6. Ošlejšek, R., Sochor, J.: A Flexible, Low-Level Scene Graph Traversal with Explorers. In *Spring Conference on Computer Graphics 2005*, Bratislava, 2005, pp 194–201.
7. Pawlak, Z.: Information Systems. ICS PAS Reports, Warszawa, 1978.
8. Reiners D.: A Flexible and Extensible Traversal Framework for Scenegraph Systems. In *OpenSG Symposium*, 2002.
9. Rumbaugh, J., Jacobson I., Booch G.: The Unified Modeling Language Reference Manual, Second Edition. Addison-Wesley, 2005.
10. Gamma, E., Helm, R., Johnson, R., Vlissides J.: Design Patterns Elements of Reusable Object-Oriented Software. Addison-Wesley, 1995.
11. Zezula, P., Amato, G., Dohnal, V.,Batko, M.: Similarity Search: The Metric Space Approach. 2005. New York, Springer, Advances in Database Systems, Vol. 32, 2005.

Automatic Annotation of Geographic Maps

Mirko Horstmann[1], Wilko Heuten[2], Andrea Miene[1], and Susanne Boll[3]

[1] Center for Computing Technologies (TZI),
Universität Bremen, Postfach 330440,
28334 Bremen, Germany
{mir, miene}@tzi.de
http://www.tzi.de/bv
[2] OFFIS
Escherweg 2
26121 Oldenburg, Germany
heuten@offis.de
http://www.offis.de
[3] University of Oldenburg
Department of Computing Science
Escherweg 2
26121 Oldenburg, Germany
Susanne.Boll@informatik.uni-oldenburg.de

Abstract. In this paper, we describe an approach to generate semantic descriptions of entities in city maps so that they can be presented through accessible interfaces. The solution we present processes bitmap images containing city map excerpts. Regions of interest in these images are extracted automatically based on colour information and subsequently their geometric properties are determined. The result of this process is a structured description of these regions based on the Geography Markup Language (GML), an XML based format for the description of GIS data. This description can later serve as an input to innovative presentations of spatial structures using haptic and auditory interfaces.

1 Introduction

Many of our daily tasks require knowledge about the layout and organisational structure of physical environments. These tasks include navigation and orientation as well as communication about geographic locations and are most often supported by maps. A map can be seen as a two-dimensional representation of a real world environment with a reduced amount of information that is created for a specific context and goal. Information stored in maps can be used to build a mental model of the physical space and to understand geographical entities and relationships.

Maps use graphical representations to visualise an areas' spatial layout and the semantic entities it contains such as parks, gardens, buildings and streets. Most of the existing map material is stored and managed in Geographic Information Systems (GIS) by the publisher and typically includes the modelling of the geographic area and the different elements and layers that belong to it. However, in many cases an end user only

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 69–76, 2006.

receives a bitmap of the map in which all semantic entities only exist implicitly, encoded as coloured pixels. This almost complete loss of semantic information excludes many people from using the geographic maps as an orientation and exploration support: Visually impaired and blind people cannot see the layout of the map, illiterate people do not have access to the text included in maps, people with motor deficiencies cannot point to tiny elements on the map as there is no option to zoom into them in a semantic fisheye fashion. Unfortunately, visually impaired people rely even more on the information stored in maps than sighted people because building a detailed mental model is required as a preparation for tasks like navigation and orientation in unfamiliar environments.

A number of techniques can be employed to make maps accessible. The most widely used method are tactile printouts of the maps, which can be produced on swell paper or, in a more complex process, with thermoforming.

However, tactile diagrams suffer from their relatively low resolution and the limited ability of the finger tips to recognise fine structure. Patterns that symbolise different areas on a map are therefore limited to a few distinguishable textures. Furthermore, lines of Braille text are usually 6mm high and cannot be reduced in size and thereby further clutter the tactile image. The result of these problems is that existing maps usually have to be completely redesigned in order to present them in a tactile format. There is no fully automatic process for a conversion as one might think. Nevertheless, tactile maps are an important aid for blind people to make themselves familiar with new environments, as is shown in [6].

Projects like TACIS [1] or, more recently, BATS [2] have tried to overcome the limitations of tactile maps with combined tactile and auditory output. Like most approaches, which use specialised, non-visual presentations to convey information, they rely on the conversion of existing material, which often means that a laborious manual process must be applied if this material is not already in a structured format that includes semantic descriptions. The TeDUB [3] system therefore aimed to semi-automatically interpret simple diagrams for an accessible presentation using image and knowledge processing techniques.

Our approach to the problem is a software system that extracts semantic entities from maps provided as bitmap images. The software first identifies coherent regions of similar colour and classifies them as one of several known types. Next, nearby regions are grouped to form single entities. For each of these, a structured representation of its shape and its type is generated in the standardised Geography Markup Language (GML). This description scheme then forms the semantic annotation of the map, which can be used in various ways, e.g., for non-visual representations like haptic, tactile or auditory display as well as any multimodal combination. With this information the user can build a mental model and familiarise with a physical environment. The proposed use-case is that of exploration of a given area for an overview, rather than for exact navigation. This first implementation is therefore meant as an alternative to tactile orientation maps. Our proposed solution focuses on city maps. However, most ideas can be applied to other map types. Maps may later enter the system through a scanner interface for printed material or as bitmap images from web pages.

2 Requirements for the Software

In this section, we identify the requirements for a system that extracts semantic information of existing city maps automatically and provides this information in a format that can be used by other systems for a non-visual representation for blind and visually impaired or otherwise print impaired people. The analysis addresses three topics: The kind of information that has to be conveyed to the user, the requirements for existing maps from which to extract the information, and the interchange format for the information.

2.1 Entities of City Maps

There are no known open standards as for what objects a city map should consist of, publishers tend to establish their own standards. Therefore we have compared several city maps of various online and print publishers and have found the following set of typical objects:

– Parks and gardens
– Water (lakes, rivers, seas)
– Streets of various types
– Squares and Places
– Quarters and other organisational structures
– Monuments, sights, points of general interest (hotels, shops, . . .)
– Public buildings (churches, schools, town hall, . . .)
– Public transportation information (stations, routes, . . .)
– Bridges and tunnels
– Additional objects to help interpreting the map (keys, scale, north arrow, . . .)

Most of these items are associated with additional attributes like their names or one-way directions for the traffic. Moreover, a city map implies the location of objects relatively to other objects based on a mapping from the real world. The city map mediates shapes of certain objects-types. With the help of a scale a user can measure distances and determine sizes of objects.

In order to familiarise with a city and to get an initial overview, large objects, their shape and location as well as distances between objects are more important than small streets and single buildings. Furthermore, too much information at the same time will make it more difficult for the user to build a mental model of the depicted city. Therefore, the presentation should only include larger objects or groups of smaller objects of the same type, which are geographically close to each other. E.g., single houses should be grouped to blocks, several blocks to residential areas.

In order to reduce the amount of information, which is presented at the same time, functions that appear in current map viewers are also useful for non-visual representations. These include filtering objects, changing level of details, zooming and panning. These have to be implemented in the specific viewer and are not covered in this paper, although the underlying format used to communicate the data must support these functions (see Section 2.3) .

2.2 Maps for Semantic Extraction

Maps usually come in two formats: bitmap images and vector graphics (GIS data). Although vector graphics are more amenable for the task of extracting the necessary spatial information about areas of interest, it is often the case that bitmap versions are more easily available (e.g., on web pages or as scanned images from printed maps) whereas vector version usually have to be obtained commercially and are then rather restricted regarding their use and redistribution.

2.3 General Requirements for Modelling Semantic Annotations of City Maps

It is important that the format for storing the extracted semantic information is open, easy to read and easy to distribute. Therefore a standardised modelling language for geographic entities is strongly recommended. In order to share semantic information between other maps and publishers, the semantic information should be stored separately from the map itself. The format should be powerful enough to describe the entities listed in Section 2.1 and their attributes. Keeping the later non-visual representation in mind, the description of the geographic objects should not be in a visual format, e.g. pixels, but rather in a vector format, which can be used for haptic and auditory rendering. Furthermore a publisher should be able to extend the description for individual needs.

3 Automatic Extraction of Semantic Information from City Maps

There are various types of maps which code geographic information in different ways. Nevertheless, sighed people are able to identify the most important things at a glance. In this paper we focus on city maps with a typical set of colours to distinguish different entities like watercourses, parks or buildings and our approach makes use of this special kind of colour code for an automatic extraction of entities. After a pre-processing step that removes text and noise from the image, image regions with a specific colour (one that is within a given interval of values for the separate colour channels) are detected. This results in a set of image masks of which each one marks all regions of one respective kind of entity. Nearby regions are then grouped to form areas with the respective entities.

Since we do not consider textual information printed on the map in this first prototype, our first processing step is to remove text through morphologic operations. An example is shown in fig. 1 where black text is removed through a morphologic closing operation. This approach is simple but suitable to remove text on most maps. If a map includes text in different colours, our more advanced approaches for text/background separation could later be employed (see e.g. [5,4]).

The next step is to find image regions with a specific colour which represent certain objects. The colour of each requested object is specified by colour intervals given by a minimum and maximum value for each of the red, green and blue colour channels. To reduce noise the image is smoothed with a median filter. To segment the image regions belonging to a known entity the image is binarised using the colour intervals specified for that entity. The resulting image mask shows pixels that are within the given colour intervals and therefore represent the requested objects. Small areas which are irrelevant for further interpretation are removed from the binary image by morphologic operations. Figure 2 shows an example for the segmentation of watercourses.

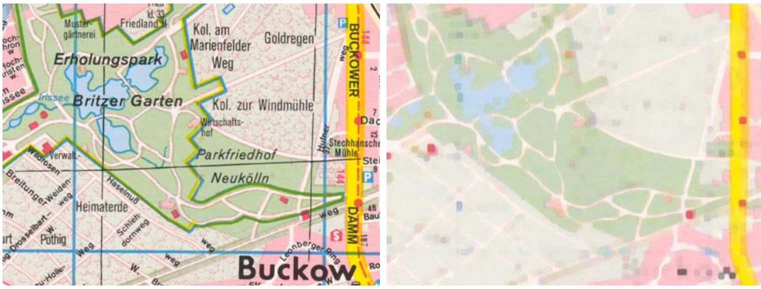


Fig. 1. City map example "Britzer Garten" (original, left) and after removing textual information (right) © Falk Verlag, http://www.falk.de

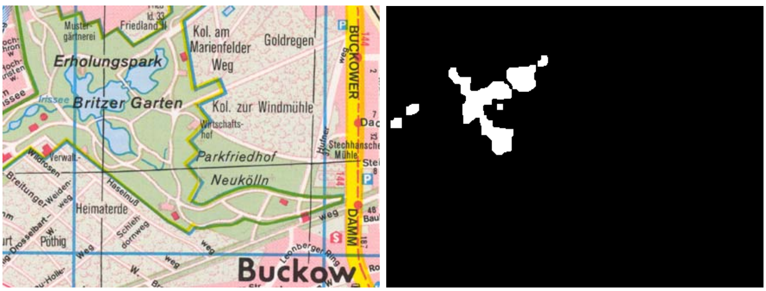


Fig. 2. City map example "Britzer Garten" and segmentation results for watercourses © Falk Verlag, http://www.falk.de

Fig. 3. Clustering of watercourses with a distance threshold of 15 (left) and 60 (right) © Falk Verlag, http://www.falk.de

Often, objects are split into several image regions by other objects. A typical example is a park which is split into several green image regions by paths or roads crossing through it. Such regions have to be grouped together and treated as one object during further interpretation. To group these regions together a threshold is specified up to which distance regions of the same colour are clustered together. A cluster of regions is then represented by its convex hull which is later described as a polygon. The clustering step also allows us to group together several buildings to a building complex.

Fig. 3 shows how the clustering process is influenced by the distance threshold.

4 GML for Modelling Semantic Information on City Maps

The Geography Markup Language (GML) – an initiative by the Open Geospatial Consortium (OGC) – enables the specification of two- and three-dimensional geographical objects (also referred to as features). It provides only a *general* framework for describing geographic features (e.g., it provides a “polygon” element rather than “lake”, “forest” or “building” elements). It is the application developer’s task to specify his or her own application schema. By using and extending the GML we can describe and model the extracted semantic information discussed in Section 2.1 and ensure that the format is open and standardised and therefore readable by everyone who wants to convey geographical information to blind and visually impaired users.

Section 2.1 lists a set of geographic entities that we need to describe – a “collection of features” in the GML nomenclature. GML provides abstract types for features as well as collections. In order to use them, they must be extended to form concrete types. For our schema, we have therefore defined two element types “GeographicFeatureCollectionType” and “GeographicFeatureType” based on the mentioned abstract types. From these types, we instantiate two elements “FeatureCollection” and “Feature”, the goal of the FeatureCollection element is to hold all Feature elements:

```
<xs:element name="Feature" type="en:GeographicFeatureType"
  substitutionGroup="gml:_Feature"
/>

<!-- root element --> <xs:element name="FeatureCollection"
   type="en:GeographicFeatureCollectionType"
   substitutionGroup="gml:_FeatureCollection"
/>
```

Our features are either buildings, parks, lakes, sights, or squares. Therefore, we have defined a required additional attribute to the “GeographicFeatureType”, which forces the GML instance author to categorize a feature as such:

```
<xs:simpleType name="GeographicDescriptionType">
   <xs:restriction base="xs:string">
      <xs:enumeration value="Building"/>
      <xs:enumeration value="Park"/>
      <xs:enumeration value="Lake"/>
      <xs:enumeration value="Sight"/>
      <xs:enumeration value="Square"/>
       ...
   </xs:restriction>
</xs:simpleType>
```

There are several standard object properties that we can assign to each feature. For example, we can assign a name and a description to a feature. However, this is not mandatory. We can also specify the bounding box of the feature if we wish. For our purpose, however, the really important information is the definition of the features as polygons and their vertex coordinates as well as their “featureType”.

The GML location element allows us to specify the location of a feature as a polygon element, whereby the vertex coordinates are given. Features can also be points, curves, multi-point lines, as well as more general objects. Therefore by using this schema, we are able to add new geographic elements easily. In the gml:Polygon element, we can specify the exterior using the “gml:exterior” element. A polygon exterior consists of an exterior linear ring. The coordinate pairs of this ring are the vertex coordinates of the polygon. Thus, there must be at least three vertices, with the last pair of coordinates being the same as the first. The coordinate pairs are also referred to as control points of the linear ring. We can specify these control points with the “gml:pos” element, whereby we enter the coordinate pair separated by a whitespace. For example the description of a building would look like this:

```
<en:Feature featureType="Building">
   <gml:location>
      <gml:Polygon>
         <gml:exterior>
            <gml:LinearRing>
               <gml:pos>619 209</gml:pos>
               <gml:pos>643 125</gml:pos>
               <gml:pos>706 84</gml:pos>
               <gml:pos>716 99</gml:pos>
               <gml:pos>677 228</gml:pos>
               <gml:pos>619 209</gml:pos>
            </gml:LinearRing>
         </gml:exterior>
       </gml:Polygon>
    </gml:location>
</en:Feature>
```

In this case, the coordinate pairs are simply the pixel coordinates of the bitmap image. GML provides support for various coordinate systems, so that instead of these pixel coordinates we could later use Gauss Krüger coordinates.

5 Results and Future Work

In this paper we present a promising approach for improving the accessibility of maps. We show that automatic methods can be applied to extract geographic information from maps, which leads to a semantic annotation. We are using the open standard GML to describe the extracted semantic information in a structure, which can be transformed to auditory or haptic presentations. Our solution demonstrates a high potential in helping people with special needs to access maps in bitmap format, which is the most used format on the Web.

Our future work will concentrate on two aspects: First, the exploitation of other image features like text or symbols to extract additional kinds of information. This will not only extend the amount of useful information extracted from city maps but it will also allow us to investigate the usefulness of our approach for other kinds of maps that do not as much rely on colour codes. Experience in image processing shows that it is not to be expected that more complex information can always be extracted automatically, so a semi-automatic process should be aimed at. Second, a closer integration of the methods into an existing workflow. This will allow a content creator to make use of the methods in the larger scope of accessible map creation.

Acknowledgement

This paper is supported by the European Community's Sixth Framework Programme, FP6-2003-IST-2-004778. We would like to thank Alexander Köhn for his support during the development of the software prototype. The prototype uses OpenCV, the Open Source Computer Vision Library (`http://www.intel.com/technology/computing/opencv/index.htm`).

References

1. Gallagher, B., Frasch, W.: Tactile acoustic computer interaction system (TACIS): A new type of graphic access for the blind. Technology for Inclusive Design and Equality Improving the Quality of Life for the European Citizen, in Proceedings of the 3rd TIDE Congress, Helsinki, June 1998.
2. Parente, P., Bishop, G.: BATS: The Blind Audio Tactile Mapping System. In Proceedings of the ACM Southeast Conference (ACMSE '03), Savannah GA, March 2003.
3. Horstmann, M., Lorenz, M., Watkowski, A., Ioannidis, G., Herzog, O., King, A., Evans, D. G., Hagen, C., Schlieder, C., Burn, A.-M., King, N., Petrie, H., Dijkstra, S., Crombie, D.: Automated interpretation and accessible presentation of technical diagrams for blind people. In New Review of Hypermedia and Multimedia, Volume 10(2), pp. 141–163, December 2004.
4. Becker, H.: Automatische Extraktion von Szenentext. Diploma thesis, Universität Bremen, 2005.
5. Miene, A., Hermes, T., Ioannidis, G.: Extracting Textual Inserts from Digital Videos. In Proc. of the Sixth International Conference on Document Analysis and Recognition (ICDAR'01), Seattle, Washington, USA, September 10–13, 2001, pp. 1079–1083, IEEE Computer Society, 2001.
6. Espinosa, M. A., Ungar, S., Ochaíta, E., Blades, M., Spencer, C.: Comparing methods for introducing blind and visually impaired people to unfamiliar urban environments. In Journal of Environmental Psychology, 18, pp. 277-287, 1998.

Towards the Use of Ontologies for Improving User Interaction for People with Special Needs

Shuaib Karim and A. Min Tjoa

Vienna University of Technology
{skarim, amin}@ifs.tuwien.ac.at

Abstract. Formal description of concepts so that it may be processed by computers has great promises for people with special needs. By making use of ontologies, improved user interaction with personal information management system is possible for these people. An ontology using semantic web technology is proposed which formally describes the mapping information about user's impairments, and the available interface characteristics. Effort is made to enhance accessibility at a generic level by making it possible to enrich the ontology for a diverse range of users. Consequently users with all types of special needs are able to get already customized interfaces. Especially, the possible adaptation to our prototype Personal Information Management System SemanticLIFE [1] is the trigger for this investigation.

1 Introduction

The end user can conveniently perform the tasks through UI (User Interface) provided mental models of both the designer and the user, and the implemented system image are coherent with each other [2]. Practically, this triangle is never perfect due to lack of shared semantics at each point. The situation is aggravated in case of users with special needs. Accessibility for these people can be improved in many ways, e.g., improving physical access (mobility impaired people using wheel-chairs require information on accessible hotel rooms, lifts, ramps), sensory access (hearing or visually impaired people need tactile markings or hearing augmentation) & communication access (people having difficulties with written text, vision, speech, and language) [3]. Ontologies are one of the key enabling technologies for exploring the information in numerous possible ways [4]. Their exploitation for accessibility would accelerate the inclusion & improve the eQuality for people with special needs in the digital world. It is possible to formally specify the concepts in ontologies & hence remove the ambiguities so that the interaction is made conformant with the needs of the task at hand, & user abilities. The output is presented using additional devices & techniques called the Assistive Technology like screen reader, electronic Braille, screen magnifier, usage of Alt text, sticky keys etc according to the Web Content Accessibility Guidelines (see WAI[1]). It is noticeable that these guidelines mainly focus on vision impairment & to some extent on motor impairments. Even for visually impaired people mostly total blindness is assumed, whereas, there are many types of vision problems, e.g., low vision, color blindness, problem in which the person can not recognize the images even his own. Other impairments are not significantly addressed by these guidelines [5], e.g., person with memory problems, mobility impairments etc.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 77 – 84, 2006.

The implementation of existing guidelines, work well for visually impaired people provided the information is well structured. For static textual information contents on web pages this strategy is already practiced but still there are unresolved issues. Currently the mapping of impairments with accessibility guidelines is hard coded either directly in HTML documents or in style sheets. The problem is severe where one has to navigate the ever changing information like on the web & explore the hyperlinks with possibility of each time going into a different navigational structure. The issue is similar in our system SemanticLIFE which is managing the information using ontologies for making associations of concepts. These associations represent the additional knowledge which was not present before. The human users had to read the contents & make sense of it themselves. Now all users, especially the people with special needs are in a better position to access this added explicit knowledge. The accessibility guidelines will be fed in the ontology along with the impairments and interface data. The link between them is automatically managed by the ontology. The ontology management is made under user control, be it the user himself or his / her caregiver.

In the next section, related work is discussed. Next, the accessibility requirements for our system are briefly described, followed by our suggested approach for building "Impairment-User Interface" ontology. Then some important consequences of our ontology, and conclusions and future work are described.

2 Related Works

Haystack project [6] has successfully used RDF for designing a component architecture that provides rich & uniform UI. It consists of four parts, i.e., layout, informative, decorative, and view. The UI components are populated based upon the ontology contents, the metadata for whom is stored in RDF. The semantic information items to be presented are associated with appropriate view part in RDF. Here we can incorporating our accessibility ontology while implementing components.

Ontologies are used in [7] for providing accessibility in context of web page annotations for conveniently navigating the visual structure of web pages. Their ontology consists of semantics about mobility (travel objects such as way points, orientation points, & travel assistants), authoring (header, logo, label, heading, footnote, section) & context (information seeking, surveying, orientation, navigation, browsing). The WAI[1] provides WCAG guidelines to encode some of the above information semantics but not all. Also, the accessibility is provided only for visually impaired users.

The recent work by [8] for increasing accessibility of web pages, suggests an approach to encode the semantic information of the page directly into the page itself by introducing lightweight markup, without compromising the creative activity of authors and designers. An ontology is created representing the meaning of data in XHTML metatags and then encoding this meaning into the data. The relevant CSS remains unaffected, while the semantics become implicit part of data. It works for pages with available CSS. Only simple instantiations with property assertions is possible. Usage of metatags reserved for other purposes is not a stable solution. But it highlights the possible XHTML enhancement with new tags for including semantics.

---

[1] Web Accessibility Initiative of the W3C `http://www.w3.org/WAI/` (4th Mar'06).

Different kinds of user policies (navigation, exploration, presentation, etc.) are introduced in [9] for making the presentation & interaction of XML contents accessible. The user policies are supposed to take into account the user's physical & cognitive impairments by encoding the relevant information in XML style sheets. This is a nice and practical approach for keeping separate the original document contents & the user preferences. This can be combined with Haystack's design to improve accessibility.

An organized set of steps is described in [10] for designing ontological-driven semantic applications such as our prototype system SemanticLIFE. An additional step could be to incorporate accessibility using ontologies.

The WWAAC[2] project by the European Union is trying to integrate the assistive technology, the web, and the signs and symbol language used traditionally by people with communication problems for text interpretation. The accessible UI complexity is highlighted as a challenging issue to tackle due to conflicting needs by different users. Also, ontologies are used to exchange semantics between the concepts databases of symbols and the assistive technology. The framework is composed of Concept Code Definitions (a plain concepts list), Base Reference Ontology (concepts mapped from WordNet[3] to used symbols) and (Complementary Reference Ontology (specifies missing concepts). Different user groups have been successfully tested against this framework in [11]. The proprietary ontologies like Assistive technology Ontology need to communicate with the framework by mapping against Reference, and Complementary ontologies using APIs. The role of user impairments is fixed. Our impairments ontology can be used for associating concepts from Concept Code Definitions to specific impairments so that presentations are customized at run time.

The human disease is conceptualized around Type (disorder types), Symptom (signs/indicators), Cause (genetic/environmental), and Treatments (surgery, drug therapy, physiotherapy, etc.) in [12]. The therapy can be extended to incorporate the notion of interface therapy / adaptation for convenience and rehabilitation of impaired user. The concepts of our interest are *type* (impairment type with associated properties to determine the severity of the impairment), and the *treatment* in context of UI.

3 Accessibility Requirements

In order to make an accessible UI, one has to list down nature of the system, the user group, nature of data, tasks users need to perform, and their ability or disability to perform their tasks. Although this list is not exhaustive, however it serves our purpose for demonstrating the positive impact of ontologies on the UI.

The input to our prototype system is user's lifetime information items captured by independent data feed modules. Those range from emails, documents of various types, calendar data, contacts, process state data, browsed web pages, chat sessions etc. The *primary goal* of our visualization is to explore the associations of information items. The *secondary goal* is to visualize their contents to explore further associations. *The associations are of three types*. Firstly, these are *based upon the already available metadata* of an information item, for example Exif header fields for

---

[2] World Wide Augmentative & Alternative Communication http://www.wwaac.org/(4th Mar'06).

[3] `http://wordnet.princeton.edu/` (4th Mar'06).

pictures, and email header information. Secondly, the *annotations given by the user*, for example email1 is related with file1. Thirdly, *the associations discovered from the contents* of the information items, e.g., finding the string "eclipse svn problems" in an email body, and associating it with web page "http://subclipse.tigris.org/".

Users of the system may vary from normal user to a user having some special needs. Since it is not a collaborative but a personal system, therefore a personalized UI would fulfill the user's requirements better than a universal interface. However, using appropriate style sheet the UI can be adjusted for any user. The UI should be conformant with the accessibility checklist mentioned in WAI[1].

4 Suggested Approach

The mapping between the information about user's impairments & available interface characteristics is formally described in an ontology. Then simple subsumption mechanism can be applied based on class-subclass hierarchy & sophisticated inference mechanisms by description logic statements for automatic interface adaptation.

According to our requirements we have followed suitable guidelines from various ontology development approaches [13] [14]. The activity is divided in two steps, i.e., finalizing general ontology characteristics, and designing its contents. However the ontology development, especially the content design phase is to be iterative in nature.

General Characteristics. It is important to know the purpose, and scope of the ontology. Then its possible interaction with other ontologies is specified, followed by a description of users, some motivating scenarios, and a list of competency questions.

Purpose and Scope. The ontology is about user's physical and cognitive impairments, and the related consequences on the UI. It will be used in context of user's personal information management system. The objective is to exploit the impairments data so that user's interaction with the system is optimized in terms of improved UI. *The aim of this ontology is to make improvements for all the impaired users in a generic way instead of focusing only on the stereotypical cases of impaired users.* By exploiting this generic ontology, improved personalized interaction for all could be possible to accommodate diversity. It is a domain specific ontology and will not cover general world or common sense knowledge in the initial attempt.

Interaction with Other Ontologies. In our current prototype, the impairments ontology will be required to interact with existing UI ontologies. Other possible interactions could be with DOID[4], UMLS[5], ICD9CM[6], and MeSH[7]. Sample extracts from DOID and MeSH are:

---

[4] `http://obo.sourceforge.net/cgi-bin/detail.cgi?disease_ontology` (4th Mar'06)

[5] `http://www.nlm.nih.gov/research/umls/` (4th Mar'06)

[6] `http://icd9cm.chrisendres.com/` (4th Mar'06)

[7] `http://cvs.sourceforge.net/viewcvs.py/obo/obo/ontology/vocabularies/mesh.obo` (4th Mar'06)

[Term] id: Doid: 3203, name: Blindness of both eyes, impairment level not further specified, rank: 7, xref_analog: Umls_Cui: C0271217, xref_analog: Icd9cm_2005: 369.00, xref_analog: Umls_Icd9cm_2005_Aui: A0241434, is_a: Doid: 3204

[Term] id: Mesh: A.01.456.505.420, name: eye, synonym: "eye" [], synonym: "ophthalmologic_effect" [], synonym: "ophthalmological_effect" [], is_a: Mesh: A.01.456.505 ! face, is_a: Mesh:A.09 ! sense_organ

Users and User Scenarios. The ontology will be used by the end user or his / her caregiver in case of severe impairments like cognitive problems. User interaction is not only limited to traditional UI part. This ontology is usable in many ways, e.g.,

a. The ontology will be used for customizing / adapting the UI according to user impairments. Each interface component has a certain affordance of use [15]. If the user is impaired to use it in that way, then it should not be part of the UI. Instead a component with better affordability would be suggested for this user, e.g.,
   - Color blindness can be Red-Green, Blue-Yellow, and Monochromacy (complete inability to distinguish any color). The confusing colors on an interface can be avoided for a particular type of user.
   - Visual acuity refers to the clarity of one's vision, a measure of how well a person sees. Font size can be adjusted according to the user's visual acuity.
   - Blindness in one half of the visual field normally due to stroke. For this user, the information should only be presented on the better half of the screen.

b. In tourism domain, the ontology can provide accessibility for a better travel planning. But, it is not the major focus for this paper. If user is mobility impaired, then while searching route from point A to point B the availability of accessible lifts and restrooms can be shown on transits, & also the corresponding time calculations for the journey (e.g., accessibility of `http://www.vor.at/`)

Questions to Be Answered. Some competent queries are framed which must be answered by our ontology. For example, is UI component (such as vertical scrollbar) suitable for the user? Otherwise find the alternate interface component. What is suitable font size? Which colored control buttons are suitable? What is the most suitable screen area for presenting information (right, left or central)? Is lift facility available on a specific underground station? Are textual descriptions available using Braille for a certain museum etc? Finally, as part of evaluation, all the queries should be satisfied and should give correct, consistent, and reliable results without any regression.

Design of Ontology Contents. Three approaches for designing the ontology contents are mentioned in [14], i.e., top-down, bottom-up, & mixed. We followed the top-down approach which calls for collection of terms, finding/making hierarchy of terms, & finding synonyms. This is followed by identification of concepts, their attributes, restrictions, cardinality, & interrelations of concepts.

Important Terms and Their Organization. The intention of this activity was to find the concept hierarchies & their interconnections. We also interviewed some physicians[8] for this activity. A sample taxonomy is made & is available in technical report[9].

---

[8] Private communications with Zubair Kareem, M.D. Phy-Neurologist (Holyoke, MA).

[9] Impairments Taxonomy `http://storm.ifs.tuwien.ac.at/publications/ImpTax.pdf` (2nd May'06).

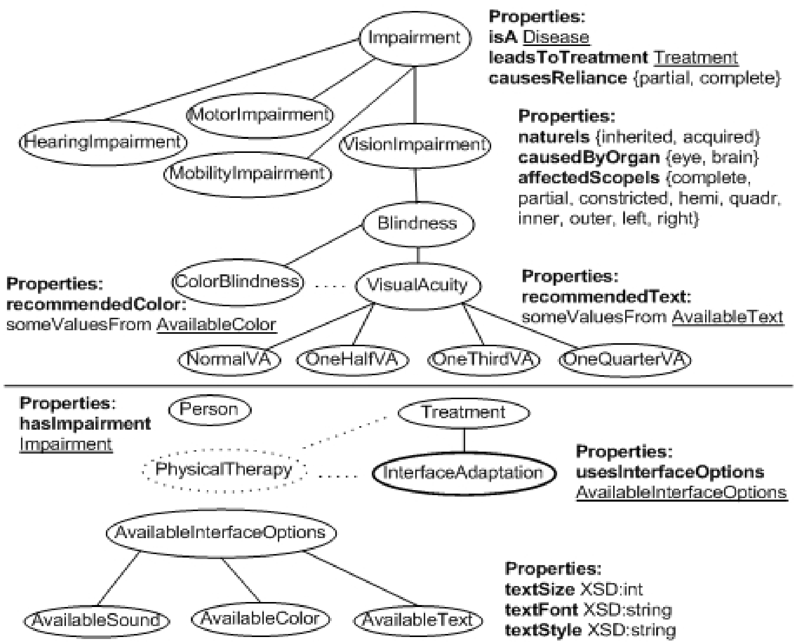


Fig. 1. Part of Impairment-User Interface ontology

Classes and the Class Hierarchy. Extending the work from [12] an ontology is created using OWL-DL in Protégé 3.1. *The main extension is the InterfaceAdaptation (see Fig. 1) as another treatment for impaired users of computer systems.*

"Impairment" is a type of disease which needs to be treated. As a result of impairment the concerned person can become partially or completely reliant on some one. The sub-classes of "Impairment" and other classes are managed with the help of given properties. As a sample case, visual acuity is taken care of by binding it to appropriate text sizes. The sub-classes of "VisualAcuity" are associated with "AvailableText" by specifying restriction for its property "recommendedText". While entering the user's impairments data (using instance of "Person") user's visual acuity instances would be linked with suitable text size. Then it can be used for displaying text on the interface for this user. A very simple case is chosen as a proof of concept.

5 Consequences

The consequences of this ontology are numerous. Primarily, it will be helpful in adapting the UI for a specific user, e.g., specifying a suitable text size. It can also be used in deducing the best match for a user with multiple impairments, e.g., if user has a very low visual acuity and also he can not focus on left part of the screen then it is a composition of impairment classes and related interface options. Also, the historical data generated by users would help to get an insight about the evolving cause-effect relationship between the impairments and the computer interfaces.

6 Conclusions and Future Work

Presenting information according to user's special needs is a complex task. In absence of a generic solution each type of special need results in different interface adaptation effort. The contributing factors, i.e., the impairments (medical domain), interfaces (computer and assistive technology domain), and the consequences of impairments on interface (accessibility guidelines) are brought together in form of Impairments-User Interface ontology. The demonstration by using a test case shows that the ontology could provide a generic step in improving universal accessibility.

Our future work will focus on incorporating other impairments. Available ontologies will be used after transformation, if their format is not OWL. Currently, reasoners only work on ontology classes. Later, we also plan to apply those on individuals.

Our approach & projects like USERfit [16] (generating usability specifications using assistive technology), & GADEA [17] (UI development framework adapted to human cognition diversity) may also help & complement each other in future.

Acknowledgements

The work is supported by Higher Education Commission of Pakistan, and ASEAN-EU University Network Initiative. Special thanks to Zubair Kareem, M.D. Phy-Neurologist for useful remarks on impairments & its potential use for rehabilitation.

References

1. Ahmed M., Hoang H.H., Karim M.S., Khusro S., Lanzenberger M., Latif K., Michlmayr E., Mustofa K., Nguyen H.T, Rauber A., Schatten A., Tho M.N. and Tjoa A.M., 'SemanticLIFE' - A Framework for Managing Information of A Human Lifetime, Proceedings of the International Conference on Information Integration, Web-Applications and Services, IIWAS'04, 27-29th Sep. 2004 Jakarta-Indonesia.
2. Dix A., Finlay J., Abowd G., and Beale R., 1993. Human-Computer Interaction, Prentice-Hall, New Jersey.
3. Pühretmair F., It's Time to Make eTourism Accessible, Proceedings of the 9th International Conference on Computers Helping People with Special Needs ICCHP'04, LNCS 3118 (2004), pp. 272–279.
4. Garcia E. and Sicilia M-A., User interface tactics in ontology-based information seeking. PsychNology Journal, volume 1(3) December 2003, pp 242–255.
5. Poulson D. and Nicolle C., Making the internet accessible for people with cognitive and communication impairments. International Journal Universal Access in the Information Society, volume 3(1) March 2004, pp 48 – 56.
6. Karger D.R., Bakshi K., Huynh D., Quan D. and Sinha V, Haystack: A Customizable General-Purpose Information Management Tool for End Users of Semistructured Data, In CIDR, January 2003, California-USA.
7. Plessers P., Casteleyn S., Yesilada Y., Troyer O.D., Stevens R., Harper S. and Goble C., Accessibility: A web engineering approach. Proceedings of the Fourteenth International Conference on World Wide Web, Sep, 2005, pp 353–362, Chiba-Japan.

8. Harper S. and Bechhofer S., Semantic Triage for Increased Accessibility. IBM Systems Journal issue 44-3, 2005, pp. 637-648.
9. Encelle B. and Baptiste-Jessel N., Adapting Presentation and Interaction with XML Documents to User Preferences, Proceedings of the 9th International Conference on Computers Helping People with Special Needs ICCHP'04, LNCS 3118 (2004), pp. 143–150.
10. Sheth A. and Avant D, Semantic visualization: Interfaces for exploring and exploiting ontology, knowledgebases, heterogeneous content and complex relationships, NASA Virtual Iron Bird Workshop, (2003), CA-USA.
11. Nicolle C., Osman Z., Black K. and Lysley A., Learning from Internet Requirements of People with Communication Needs, Proceedings of the 9th International Conference on Computers Helping People with Special Needs ICCHP'04, LNCS 3118 (2004),pp.121-128.
12. Hadzic M. and Chang E., Ontology-based Support for Human Disease Study, Proceedings of the 38th Hawaii International Conference on System Sciences, HICSS'05.
13. Kalyanpur A., Hashmi N., Golbeck J. and Parsia B., Lifecycle of a Casual Web Ontology Development Process, The 13th International WWW'04 Conference: Workshop on Application Design, Development and Implementation Issues in the Semantic Web, (May 2004).
14. Noy N.F. and Hafner C.D., The State of the Art in Ontology Design, AI Magazine, 18(3), 1997, pp. 53-74.
15. Norman D., the Design of Everyday Things.Doubleday: New York, 1990.
16. Abascal-González J., Arrue-Recondo M., Garay-Vitoria N., Tomás-Guerra J. and Velasco-Nuñez C.A., USERfit Tool: A Design Tool Oriented towards Accessibility and Usability, UPGRADE The European Journal for the Informatics Professional – Human-Computer Interaction: Overcoming Barriers, vol. IV, issue 1, 2003, pp. 5-11.
17. González-Rodríguez M., Moral-Pérez E.D., Paule-Ruiz M.D.P. and Pérez-Pérez J-R, GADEA: a Framework for the Development of User Interfaces Adapted to Human Cognition Diversity, UPGRADE The European Journal for the Informatics Professional – Human-Computer Interaction: Overcoming Barriers, vol. IV, issue 1, 2003, pp. 27-30.

A Model for Accessible Information Networks – Findings of the EUAIN Project

Benoît Guillon[1], Catherine Desbuquois[1,2], and Dominique Burger[1,3]

[1] Association BrailleNet
[2] Ministère de la Culture et de la Communication
[3] INSERM U592
{benoit.guillon, catherine.desbuquois, dominique.burger}@snv.jussieu.fr

Abstract. This paper proposes a model for Accessible Information Networks based on the experience of the "Serveur Hélène" project and on-going research within the European Specific Support Action EUAIN. It outlines the main components of the model, that should support contractual policies with publishers, as well as technical tools for improving the accessibility of the files provided.

1 Introduction

Since 2000, BrailleNet has developed a collaborative Web server to gather files provided by various players and to produce adapted documents for the visually impaired [1]. The Serveur Hélène[1] provides organisations producing adapted documents with files of books available from a secure repository. Since January 2006, the functionalities have been extended to a digital library for individual readers using a specific hardware device. To solve the problems of files formats multiplicity and incompatibility, BrailleNet has developed an integrated publishing chain. This chain is based on XML DTBook and can produce various adapted formats on demand.

The EUAIN project (European Accessible Information Network) aims to promote e-Inclusion as a core horizontal building block in the establishment of the Information Society by creating a European Accessible Information Network to bring together the different actors in the content creation and publishing industries around a common set of objectives relating to the provision of accessible information. The CEN Workshop on Accessible Document Processing (CEN/ISS WS/ADP) is a key activity for the EUAIN network. It examines the ways in which this convergence is helping to build consensus and create new standards and technologies for the provision of information in formats that are more accessible for everyone.

This paper suggests a model for Accessible Information Networks based on the experience of the Serveur Hélène and the on-going research done by EUAIN.

[1] http://www.serveur-helene.org

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 85 – 91, 2006.

2 Overview of the “Serveur Hélène” Project

2.1 Collaboration with Publishers

The principles of this collaboration with publishers are:

- Organisations producing adapted books are registered as partners of the project. They sign a contract with BrailleNet to ensure that downloaded files are used properly, also providing a list of authorised users.
- BrailleNet signs contracts with publishers allowing to store files of books on the server and make them available to registered partners.
- Publishers provide BrailleNet with the source files of their books.
- BrailleNet makes these source files available to associations, libraries for the blinds, schools partners of the project (more than 80 organisations).
- Partners upload files on the server if they want to share them.
- BrailleNet keeps the history of downloaded files.
- BrailleNet pays retribution to publishers when books are produced in Braille or large print (about 7% of the retail price of the book for each copy performed).

This approach has many advantages:

- It creates a trust network between publishers and organisations producing adapted documents, with a trusted intermediate (BrailleNet) responsible for the use of files.
- Publishers have a single organisation to discuss with.
- Publishers can monitor the use of their files via BrailleNet.
- Organisations producing adapted books can share the work they do for making accessible the files provided by publishers.

2.2 File Formats Harmonisation

In 2002, we observed that 25 different formats of documents had been gathered on the Serveur Hélène, used by transcribers in different ways for different purposes. Thus it appeared necessary to harmonise the formats in order to solve the problems of documents' representation, formats' flexibility, extensibility and portability. BrailleNet has setup an integrated production chain of documents. The DTBook Document Type Definition (DTD) was chosen as the central representation of documents. This DTD is part of the NISO Z39.86-2002 standard, also known as DAISY 3.0. [2]

To produce DTBook document, BrailleNet has created a MS Word macro to help users in producing well structured RTF documents. Such documents can contain information about the hierarchical structure, as well as original page numbers, MathML formulas and metadata. Once uploaded on the Serveur Hélène, RTF documents are converted to XML using upCast[2] and then to XML DTBook using a XSL stylesheet developed by BrailleNet.

[2] http://www.infinity-loop.de

DTBook files can be used for producing:

- accessible HTML documents using customized navigation features (several HTML files linked one to each other using table of contents, navigation links on top and bottom of each file to reach previous or next section), link to particular CSS stylesheets and most of accessibility guideleines of the W3C/WAI.
- Duxbury files (dollars-code format). Resulting files can be imported directly into DBT and then embossed in Braille.
- PDF for large print which may be more customised according to user's preferences later on (for instance images size and colors …)
- Text only Daisy 3.0 package: an ncx file for navigation and a SMIL files for synchronization are automatically generated from a dtbook file. The dtbook file is modified to include references to SMIL files. Used with an opf file including book's metadata generated by the Helene Server, a full text-only Daisy 3.0 package is built. This package can be read with a Daisy 3.0 player.

We developed XSL stylesheets to convert dtbook files in these different formats. These stylesheets are launched on the fly upon request of the user. The user can customise converters according to users needs (font size for large print, contraction or not for Braille, CSS stylesheets to be linked to HTML documents …).

2.3 Individual Secured Reading

The Serveur Hélène was primarily intended to partners who produce adapted paper books. But once the critical number of 2.000 titles was reached, and many demands from individual readers had beeb received, it became obvious that the service should be extended to an digital library for the visually impaired, and deliver e-books to any reader, interested in any title, available in a proper format. This delivery scheme offers more flexibility than the current distribution channels (libraries for the blind, associations, schools), because it is highly available and does not imply a delay in the content reception. Structured electronic documents promise a new reading experience for visually impaired enhancing access to information and rendering customisation.

We set up a protocol to secure the dialog between an hardware platform (like a Braille displays or e-book device) and the Serveur Hélène [3]. It is assumed that the platform has direct access to the Internet and can embbed a dedicated software reader. The platform prevents unauthorised exportations while documents are read. The protocol handles platforms' registration, users' subscription, rights management, catalog interrogation and books' delivery. The firts implementation has been made by Eurobraille in their IRIS platform.

3 Observations Made During the "Serveur Hélène" Project

3.1 Agreements with Publishers

In April 2006, 2.522 titles were available in the catalogue: 1.937 titles under copyright (76.8%) and 585 titles from public domain (23.2%). Since its creation, the Serveur Hélène received 8.210 files of books.

1.719 of the 1.937 titles under copyright were authorised by publishers (88.7%), authorisations pending for for 114 titles. 104 titles were refused (5.3%). In general, when publishers enter in collaboration with the "Serveur Hélène", they usually give their authorisation for every title they are demanded.

3.2 Source Files Provision

66% of the books authorised by publishers were provided with the corresponding files. For 34% of books the files were provided by partners (adapted formats) . Only 4% of titles files were missing the files. 47% of the files provided by publishers were in Microsoft Word format or RTF. 42% were provided in PDF and about 10% in Quark XPress format. The last 1% was provided in HTML.

The Case of PDF Files. As PDF is becoming a format more and more often used by publishers, we conducted a study on 196 files provided by different publishers between 2003 and 2006, in order to better identify their features. The PDF files were scanned using a tool called "pdfinfo", a tool created by the developers of Xpdf, a free open source PDF reader[3]. For each inspected file, pdfinfo prints out the version number of the PDF format, if it is tagged, optimized, encrypted, the name of the software programme used to produce this file, the name of the software programme used to create the content of the file. The results are summarised in table 1.

Table 1. PDF versions of files received from publishers

PDF version number	*Number of files*	*Percentage*
1.0	1	0.5%
1.1	1	0.5%
1.2	32	16.5%
1.3	115	58.5%
1.4	44	22.5%
1.5	3	1.5%
1.6	0	0%
Total	196	100%

The study stressed the fact that the PDF files were mostly generated with Adobe Distiller (71% of the studied PDF files), but many different versions of Distiller were used (3.0, 4.0, 4.05, 5.0, 5.01, 5.05, 6.0 and 7.0), that is why we have encountered so many different versions of PDF.

16% of the studied documents were created using Adobe InDesign, 21% using Quark XPress and 13% using Microsoft Word. It is important to note that this information was not available in 44% of the studied files.

- The majority of PDF files provided by publishers were inaccessible because of the version of PDF: Tagged PDF was first introduced in the specifications of PDF 1.4, and the support of tagged PDF was really improved with PDF 1.5. As 76% of the files were provided in a version earlier than 1.4, these files do not support accessibility features at all.

[3] http://www.foolabs.com/xpdf/

- When accessibility features were available in the PDF formats used, these features were not present in the files we received for two reasons:
  - Authoring tool or page layout software programmes were not properly used to add structure to documents;
  - Exportation tools used to create PDF files were not properly used or configured and so documents structure was present but not exported.

3.3 Statistics on the Use of the Service

2558 titles under copyright have been downloaded by registered partners since the end of 2001 (about 47 titles per month). The number of download titles has considerably increased since 2003 and currently between 200 and 300 titles are downloaded every month. Currently 64 organisations producing adapted documents are registered. These organisations are mainly located in French speaking countries and especially France.

Since January 2006, 55 readers have registered to the "Bibliothèque Hélène". 9 accounts were created by special schools for their pupils. These accounts are not bound to a particular pupil but are made available to everyone. From January to April, 2006, 44 books have been downloaded using a secured platform. Most of them were under copyright.

4 Proposed Enhancements

The results presented above validate the model of the Serveur Hélène. They show that publishers are willing to cooperate and provide files, and that there is an audience for the Serveur Hélène and the digital library.

Nevertheless, we have identified limitations on two main aspects:

- The quality of files provided by publishers that are poorly structured and re-usable;
- The fact that for now the digital library can only be accessed with one vendor specific platform.

To consolidate the implementation of this model, we propose further developments in two directions:

4.1 Instruments for Publishers and Authors

Guidance and publishing support should provided for improving the quality of their files (Figure 1):

- Authors are the first actors in the publishing process. They should therefore be involved in structuring the source documents. Publishers often provide their authors with authoring guidelines, especially in technical publications. These guidelines should include accessibility requirements and precisely describe how authoring tools should be used for that. When authoring tools do not offer the necessary features, publishers should recommend other solutions.
- As publishers often sub-contract a part or the whole pre-press process, they should include accessibility requirements in their contracts with subcontractors. This could

take the form of checklists accompanying the delivery of a file by the sub-contractor, where each item should be validated. To ease the work of pre-press technicians, specific guidelines should be published for the main authoring tools on the market. These guidelines should be written by software publishers in collaboration with accessibility experts.

- Electronic book sellers should also be provided with guidelines to make their e-documents accessible.
- Finally, organisations cooperating with publishers for the adaptation of their books should develop their expertise on the publishing software tools.

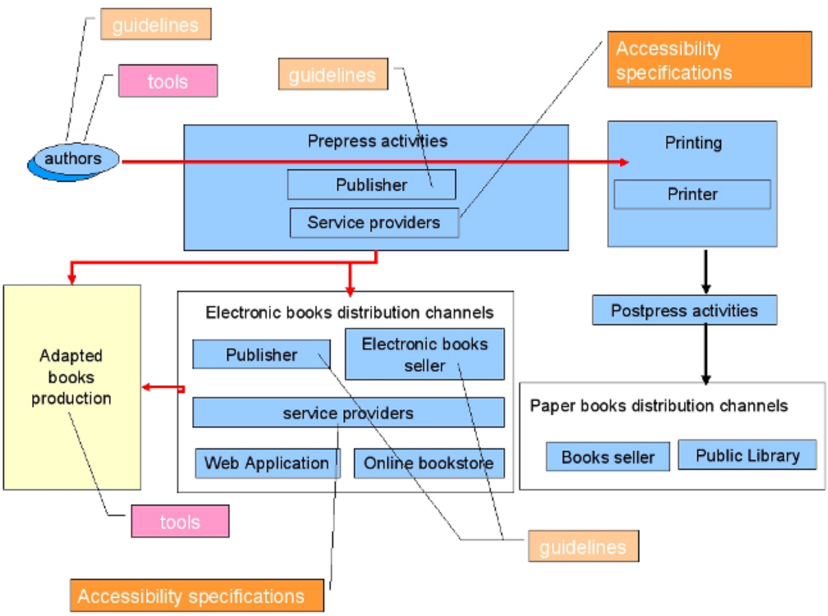


Fig. 1. The instruments enhancing the accessible information publication process

4.2 New Solutions for Distribution to Readers

This digital library should propose a simple access to documents, with no specific hardware requirements, because these requirements limit the number of potential readers. A new distribution channel must be set up, where security is still guaranteed by hardware protections (using an USB cryptographic dongle for example) but where users can access content using a simple desktop computer. The PDF format and the accessibility features of the latest version of Adobe Reader provide solutions that will be explored.

5 Conclusion

In this paper we show that it is possible to set up a practical accessible information network of organisations cooperating for the production of accessible information.

This network shall use normalised tools and practices, with technical experts to keep working with publishers for innovative solutions.

The distribution of accessible electronic documents shall reliably guarantee intellectual property rights without restricting access to information.

This network acts as a trust network where the players are well identified and are responsible and accountable for the use made of the files provided by publishers.

Relations with publishers shall be organised by trusted intermediaries to discuss intellectual property rights and structured files provision. These intermediaries are legally acknowledged by public authorities.

Publishers have clear specifications on the file formats they shall provide at their disposal. If necessary, guidelines or tools are supplied to help publishers in integrating accessibility in their production chains.

Publishers introduce accessibility in the contract they sign with subcontractors. They also guide their authors in creating structured information using the prescribed authoring tools.

When new electronic products are designed, accessibility is to be considered in the functional specifications and in design and development.

The main objective of the EUAIN Specific Support Action is to contribute to the realisation of that reachable objective, in cooperation with the publishing world.

Acknowledgements

The EUAIN project is funded by the European Commission. It is a co-ordination action (contract number 511497) co-funded by the INFSO DG of the European Commission within the RTD activities of the Thematic Priority Information Society Technologies of the 6th Framework Programme.

References

1. Helene: a collaborative server to create and securely deliver documents for the blind du Bourguet, G., Guillon, B., Burger, D.AAATE 2003, 2003-09-01, Dublin (Ireland)
2. Towards an Integrated Publishing Chain for Accessible Multimodal Documents Guillon, B., Monteiro, J.-L., Checoury, C., Archambault, D., Burger, D. ICCHP 2004, 2004-07-11, Paris (France)
3. Helene Server, a new step towards individual secured reading for visually impaired people Guillon, B., Desbuquois, C., Burger, D. AAATE 2005, Lille (France)

Making the Past a Thing of the Future: Automated Workflow for the Conversion of Printed Items into Fully Structured Digital Objects Based on Common Open Metadata Standards

Claus Gravenhorst

CCS Content Conversion Specialists GmbH, Schwanenwik 32
22087 Hamburg, Germany
Claus.Gravenhorst@ccs-gmbh.de

Abstract. This article discusses the basic need for accessible information, the METAe Project and its results as related to the digitisation of printed materials, the advantages of harvesting structural metadata during the process, an overview of the digitisation technology and workflow in the docWORKS conversion solution from CCS, and a perspective and conclusion section relating to these subjects highlighting examples of where the technology is in use today and how the information can be used in the future.

1 Introduction

Information is a basic need of the society. For many centuries innumerable archives and libraries of the world formed the largest and most reliable information centre for the learning and the curious, pupils and professors, or to put it another way … for everyone.

With the help of the Internet, current information is at nearly everyone's disposal. By comparison historical material containing valuable cultural and scientific information is not as readily accessible. Right now only a fraction of contents of the today's libraries and archives is available in digital format to the general public. Usually the original can only be accessed in printed form or microfilm/microfiche, which means search, use and distribution of the information is both time-consuming and cost-intensive.

The digitisation and conversion of printed items into electronic formats were, until recently, complex and cost-intensive. Insufficient budgets and/or resources prevented extensive transformations to digital repositories. Reliable methods for long-term security and the storage of these enormous data sets were virtually unavailable.

2 The METAe Project

As the result of the METAe project (http://meta-e.uibk.ac.at), funded by the European Commission through the 5th Framework Research Program, CCS Content Conversion Specialists GmbH, Germany developed a comprehensive software

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 92 – 95, 2006.

solution, available on the market since 2003 under the brand name docWORKS/METAe. Altogether 14 partners from European countries and the USA were involved in this research project. Through the co-operation of libraries, universities and IT companies this project represented the model of a successful co-operation for the integration of library standards into the software development process to determine the new conversion technologies needed by the library community.

docWORKS is a production tool, which offers an automated, structured conversion of printed documents into digital objects, which describe the physical and logical document structure by consistent use of international XML standards. These XML documents are to be equated concerning quality and structure with born digital documents and can be transferred to digital library systems, portals, document, content and knowledge management systems.

2.1 Creation of Structural Metadata

The main goal achieved through the project was the automatic generation of administrative, descriptive and structural metadata. In previous conventional digitisation projects these could be generated with a substantial amount of manual labour and associated costs. For these reasons only a very flat indexation and/or structuring was done in many cases.

The advantages of highly structured documents:

- As "digital original" they meet the requirements for a digital long-term storage in repositories
- With the use of XML open metadata standards, the data can be migrated and transformed to meet future requirements
- With logical structures search results are improved (chapter-, article-based) and more easily accessed (pictures, captions, footnotes, etc.)
- Continuity between digitally created and digitized documents
- Data storage is maintained separately from presentation systems therefore the data is not modified making adaptation to future systems and technologies possible

2.2 Technology Overview

The generic, rule based document analysis technology covers a wide range of different document types such as books, journals, newspapers, but also scientific documents like theses, dissertations and reports. The workflow of the conversion process has been automated and simplified to make the digitisation more cost-effective and focus on mass digitisation.

The conversion process depends on the document type and can be completed automatically, semi-automatically or manually. Interactive user interfaces are available to monitor the conversion progress, as well as the verification and correction of conversion results. For conversion a rule-based, object-oriented engine is used in connection with text recognition technology (OCR), supplemented by manual and/or semi-automatic interaction capabilities.

The conversion workflow is document and application dependent and the conditions can be varied. The goal is to make processing as efficient and automated as possible. By using client server based processing, the throughput can be scaled in such a way that it meets mass digitisation requirements. A central, server-based conversion as well as quality assurance spread over internal or external resources enables distributed production workflows.

During the conversion process physical objects such as text zones, pictures, tables, advertisements, etc. including their characteristics are determined. In addition to logical structures such as chapters, captions, author, article, etc. as well as associated metadata are determined. Text zones are converted to electronic text with integrated OCR technology. The processing of historical typefaces is supported as well.

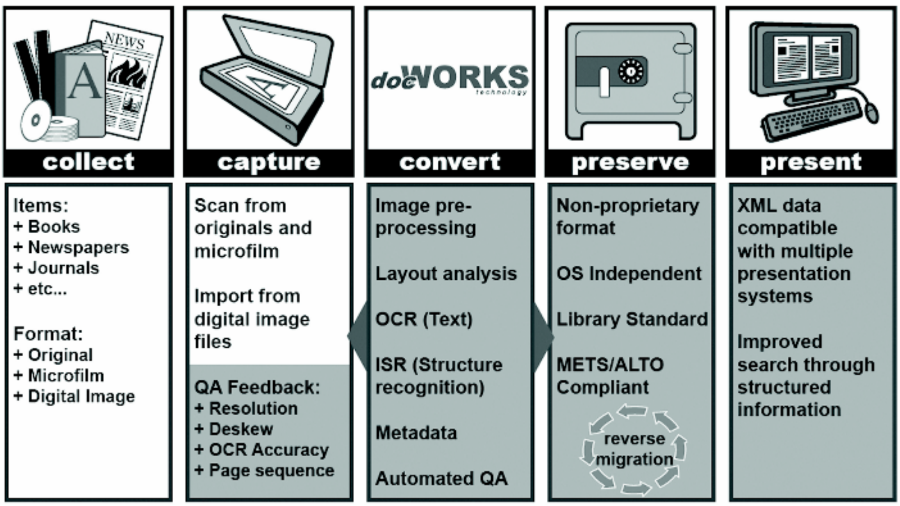


Fig. 1. Digitisation/Conversion Workflow

The rich standardized XML-based output increases the added value of digitised collections and opens up new dimensions of access and usability. docWORKS supports open metadata standards like METS, Dublin Core, MODS, NISO MIX, ALTO for storage in repositories and through XSLT transformation virtually any format can be derived for presentation and distribution purposes.

Since people with specific learning difficulties or other kind of disabilities need structured information for navigation as well as an adapted visual or audio-visual presentation via various output devices, the highly structured "digital originals" created by docWORKS provide the source for required transformations to those formats.

The converted documents are exported in different standard formats. The most important are image (e.g. tiff, JPEG, JPEG 2000), PDF (alternately with bookmarks and hidden text) and XML, where the international metadata standard METS, hosted by The Library of Congress in the USA (http://www.loc.gov/standards/mets/) is used in first place. Among other things, the "structure map" defines the logical document structures e.g. chapters and articles. In order store additional information about the physical layout from document pages, in the context of the METAe research project the ALTO schema was developed, which the Library of Congress has meanwhile adopted as the future standard for the NDNP Project (National Digital Newspaper Program, `http://www.loc.gov/ndnp/`). This project aims to digitize historical newspapers from all 50 US states and make them available online.

3 Perspectives and Conclusion

Since 2003 docWORKS has been extended to reach a higher degree of automation and cover a broader range of document types. docWORKS is in use at in-house digitisation centres at Harvard University Library, Library of Congress, Texas University Library, Royal Danish Library, National Library of Finland as well as market-oriented service vendors like OCLC Preservation Service Center in Bethlehem, USA and some others in Europe.

Although libraries, archives and universities had been the early adopters of the docWORKS technology, the publishing industry has expressed more and more interest in using effective conversion workflows for digitising their archived materials and making them available online.

The success of METAe and the subsequent projects shows that the use of new technologies can fulfil the fundamental structural requirements and form the framework for today’s digital repositories. The additional metadata generation and storage forms the basis for a digital long-term preservation as well as multiple presentation, processing and distribution options through trend-setting metadata standards with logical structures forming the basis for current and future media.

The high degree of automation fulfils the needs of mass digitisation and can thereby support and optimize existing national and international digitisation initiatives.

Accessible Navigation of Rich Media: Exposing Structure, Content and Controls in the Mobile User Interface

Markku T. Häkkinen

P.O. Box 6632, Lawrenceville, New Jersey 08648 USA
University of Jyväskylä, Jyväskylä, Finland
mhakkinen@acm.org

Abstract. As rich media moves further into our daily lives, whether at home, work, or on the go, lack of accessibility in mainstream content and end-user devices remains a challenge. Digital Talking Books are an example of how accessible content that merges text, audio, and structure, can be delivered on specialized portable devices for use by those with print disabilities. Mainstream media formats and standards are evolving that allow structural and meta information to be included in multimedia, yet authors may not utilize such capabilities and playback systems may not expose them even if present. Music playlists, as found on ubiquitous media devices such as MP3 players, offer conceptual similarities to talking book navigation but lack accessibility. With the growing presence of multimedia enabled mobile phones and devices, and increased availability of mobile audio and video, the accessibility of the content and user interface becomes a critical issue.

1 Introduction

Playlist navigation on mobile media players is an interaction model not dissimilar to that of navigating structured audio books for people with print disabilities. In fact, in most cases, both have a basic model of moving within and between hierarchically structured elements of content. In the case of music, the structure is defined by the nature of the content (Artist, Album, Song; Genre, Artist, etc.) or by the consumer ("my favorite songs," "my favorite artists," etc.). In the case of audio books, structure is defined by the producer or author, and often significantly exceeds the hierarchical depth of a music collection (e.g., Book, Volume, Chapter, section, sub-section). Audio book structures are not limited to the chapter model [1], as a variety of works and publications exist (e.g., Book, Canto, line; Play, Act, Scene, Line; magazine, letters, articles, opinions).

Navigation of digital talking books, such as those implemented in the ANSI/NISO Z39.86-2005 standard [2], is accomplished through selection of a navigation level, with movement within a selected level controlled physically in the user interface via a previous and next button. In mainstream digital media players, audio book navigation, if present, is limited to a flat list of chapter markers. The presence (or absence) of chapters on mainstream players may be indicated by a small visual tick mark on a visual timeline, with movement between chapters controlled via previous and next

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 96 – 99, 2006.

buttons. Though such navigation capabilities are discoverable non-visually, they are far from accessible and provide no overall indication of the structure.

The richness of mobile multimedia is evolving rapidly. Podcasts, growing in popularity, generally lack any navigation features, yet often consist of a sequence of segments or stories. Some podcast formats even allow for synchronization of images and text but this capability does not appear to be widely used. Video content (video podcasts) may contain a variety of content styles, ranging from cinema to university lectures. Utilization of accessibility features in video content (e.g., captions and descriptive audio tracks) appears limited, both in terms of content and implementation in the players.

Exposing navigation options can be a useful feature for many types of content and usage scenarios. Hakkinen, et al [3] have proposed generalizing the talking book navigation model to enhance usability and accessibility of structured multimedia. As content becomes more complex, exposing navigation options and control in the user interface, particularly in an accessible manner, is a design challenge when faced with limitations in screen size and physical input mechanisms. Mainstream media players generally utilize a simple user interface model with a limited number of physical input keys or touch mechanisms. Navigation is generally begun by selecting a playlist, audio book, video, or podcast. Once selected, navigation is allowed between songs/videos/podcasts and audio book chapters (if present) using a single press of a previous or next key. In addition, time-based movement within the current playing media is possible, generally by pressing and holding a previous or next key.

For rich content, simply calling navigable points a “Chapter” is inadequate, despite wide use in DVD and audio book content. This approach, though simple to implement, applies a user interface presentation and naming convention that often will not match the authored content. The ANSI/NISO Z39.86-2005 standard [2] attempts to solve this problem by providing a “resource file” to supply accessible labels for each type of structural element within a publication, thereby allowing a player user interface to appropriately render navigation options (e.g., *go to Previous or Next Canto*, *Article*, *Story*, *Line, Scene, etc.*).

Mainstream media players have exclusively focused on a visual interface, which at best augments the interaction with clicks or beeps, followed by immediate playback of the selected item. For anything other than basic player controls, non-visual operation is severely constrained. Addressing this general problem, previous research has explored using granular, time-based navigation (e.g., Roy & Schmandt [4]) to enhance usability of structured audio in a mobile device. Gestures and audio feedback have been explored to improve usability of mobile media player interfaces (e.g., Pirhonen, et al [5]). The current interaction model of digital talking books, an area with limited empirical user research and guided largely by user requirements and feedback from an active user community, suggests avenues for new research to explore enhanced usability and accessibility for mobile multimedia.

2 Prototype User Interface

A user interface model has been proposed [6] for use on personal digital assistants to navigate audio e-books via a single “wheel switch”, thus allowing non-visual access to playback control and navigation features. This model has now been implemented in prototype form for usability research, and interfaced to both a three position wheel switch as well as a circular touch pad with a center selection switch.

In this prototype, all elements of the user interface are presented both visually and in audio. Audio prompts may be presented using pre-recorded digital audio (human voice or other non-speech cues) or synthetic speech. Visual presentation is via a multi-line text LED or color flat-panel LCD display. Audio is presented via headphones. Video content, if present, will appear on the attached flat-panel display. The prototype hardware is controlled by software running on a laptop computer. Specially developed control software provides built-in, standard media player functions and also interfaces to digital talking book software [7]. The software is instrumented to record task conditions and user response times and stores this data for later analysis.

Using the digital talking book model, the adaptable, software user interface incorporates the following key features: exposure of named (labeled) levels of granularity for navigating the current content, selection of active navigation granularity, skip ahead/skip backward by current granularity, and a "*pause… where am I?*" function. The granularity levels are effectively the structural hierarchy of the current content, with the optional addition of time-based units of movement (e.g., *skip ahead 5 seconds*) and user defined navigation points (e.g., bookmarks). *See Fig. 1.*

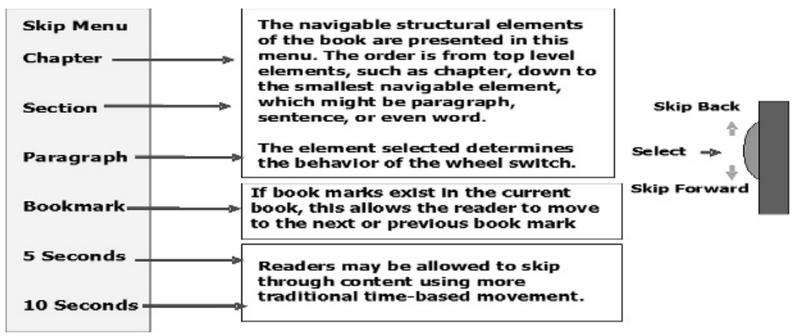


Fig. 1. Example of a Skip Granularity Menu

3 Preliminary Research

Preliminary studies are underway to explore the combined audio/visual interface with sighted and low vision users. A search and navigation task is used to examine the difference in task accuracy and completion time for both user groups. Early results show a performance advantage for the visual interface when compared to the audio only presentation. Methods to create an efficient and understandable audio interface comparable in performance to the visual model are currently being explored. This work includes the mixed use of speech and non-speech cues and variation in the speech presentation rate.

Because the hardware prototype was tethered to a laptop computer, initial studies do not place any concurrent activity load on the subject. Follow-on studies will reduce the prototype form factor to allow contextual evaluation of the interface with tasks such as walking.

4 Conclusion

Current mobile multimedia devices ignore accessibility, limiting their usability by people with visual disabilities, as well as by sighted users engaged in tasks where their eyes are “busy”. A prototype user interface has been developed and is being used to explore a combined visual/auditory presentation in the context of mainstream mobile media players. Research findings from this work are expected to have implications for the universal design of mobile multimedia interfaces.

References

1. Hakkinen, M.T. & Kerscher, G. Applying a Navigation Layer to Digital Talking Books: SMIL, XML, and NCX. Presentation at the *WWW9 Multimedia Workshop*, Amsterdam, May 2000.
2. ANSI NISO Z39.86-2005 Digital Talking Book Standard.
3. Hakkinen, M., DeMeglio, M., & Kawamura, H. gNCX: Structure-based Navigation to Enhance Usability of Digital Media. Poster presentation at WWW 2002, Honolulu.
4. Roy, D. K. and Schmandt, C. 1996. NewsComm: A Hand-Held Interface for Interactive Access to Structured Audio. In *Proceedings of the ACM Conference on Human Factors in Computing Systems(CHI’96)*. ACM, New York.
5. Pirhonen, A., Brewster, S., and Holguin, C. 2002. Gestural and audio metaphors as a means of control for mobile devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Changing Our World, Changing Ourselves* (Minneapolis, Minnesota, USA, April 20 - 25, 2002). CHI '02.
6. Hakkinen, M. Navigation User Interface for Audio eBooks. isSound Corporation Technical Presentation, August 20, 2000.
7. DeMeglio, M., Hakkinen, M., & Kawamura, H. Accessible Interface Design: Adaptive Multimedia Information System (AMIS). ICCHP 2002, Linz. Springer Lecture Notes in Computer Science.

The Essential Role of Libraries Serving Persons Who Are Blind and Print Disabled in the Information Age

George Kerscher

Secretary General, DAISY Consortium
c/o Swiss Library for the Blind and Visually Impaired, Grubenstrasse 12
8045 Zürich, Switzerland
Kerscher@montana.com

Abstract. Traditionally, libraries serving persons who are blind or print disabled have had to create accessible, specially formatted materials and distribute these products to their patrons. As new technologies evolve, there is the vision of a time when materials published for the mainstream can be made accessible to persons with disabilities at the same time and at no greater cost than the versions targeted for the mainstream consumer.

This paper and presentation will both provide you with an update on the "traditional" activities, and reveal the essential roles that many libraries for the blind are playing in the standards, open source software, and the consumer products arena that will lead to that vision of the future.

1 Libraries for the Blind Section (LBS) of International Federation of Libraries Association (IFLA) and the DAISY Consortium Working Together

LBS and the DAISY Consortium [1] have agreed to coordinate work between the two organizations. While IFLA-LBS [2] represents the interest of libraries serving persons with disabilities, the DAISY Consortium develops standards, implementation strategies and tools, and promotes technology surrounding reading and access to information. It is this simple symbiotic relationship that is proving to be highly effective. The DAISY Consortium focuses on the standards development, builds tools for production, reading, and provides training and technical support. IFLA-LBS libraries produce the accessible books and provide the surrounding services their library patrons need.

The Vision and Mission of the two organizations are different, but they share a common theme, reading, and access to information. It is this shared theme that bonds the organizations together. Without reading systems and without conforming content (books, journal articles, etc.), there would be nothing for the print-disabled library patron to read, but where do the reading systems and the content come from?

1.1 Digital Content from the Web, Digital Publishing, and Libraries Serving Persons with Disabilities

The technology savvy person with a disability has a huge advantage in accessing information today. Assistive technologies, such as screen readers for persons who are

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 100 – 105, 2006.

blind, and screen magnifiers, for persons with low vision, along with a wide range of other products that make computers accessible, also *can make* the digital information accessible. **"Can make"** must be emphasized, because the data representation of the information and the techniques used to present the data must be designed correctly to be used by everybody in our society.

A major role of libraries serving persons with disabilities is the development of the standards for fully accessible, highly functional information systems. Collectively, the libraries represent the print-disabled people they serve; the libraries are not only collections of books, but they are also centers of expertise as it relates to reading by persons with disabilities. It is natural for these libraries to work together to ensure that data models for information evolve in ways that will benefit not only the print-disabled reader, but the whole of society.

2 Delivering Services Today, Working Toward Tomorrow

The libraries serving persons with disabilities have a difficult job today. A large percentage of their patrons are not computer power users. This average library patron must be served using the technology that is appropriate for each person. Hence, the libraries are faced with the job of introducing new technologies for reading to their consumers.

But wait, there is more! The libraries are also tasked with:

- Identifying the requirements for reading systems;
- Developing the standards for data representation;
- Creating a market for hardware and software reading systems;
- Ensuring that production tools are available for creating content;
- Creating the accessible, highly functional books that conform to the standards;
- Distributing the content to the library patron.

2.1 Libraries Serving Persons with Disabilities Supporting the Formation of the Web Accessibility Initiative (WAI)

The World Wide Web Consortium (W3C) is the standards setting body for the Internet. The WAI [3] is one of the four branches of the W3C and is devoted to ensuring that the evolution of the Web includes persons with disabilities. The DAISY Consortium is a member of the W3C and is represented in the WAI activities and in other W3C working groups. Formation of the WAI was strongly supported in many sectors; libraries serving persons with disabilities helped to build that ground swell of support for the WAI. Not only were the libraries represented in the founding groups, but also libraries were driving the requirements of persons with disabilities forward in the evolving HTML, SGML, XML, and SMIL standards design.

The role of libraries in the standards setting process is essential. The requirements of persons with disabilities must be put forward, explained, designed, and *defended*. When we emphasize **defend** we mean that many, many times it is easy to design information systems that do not take into account disabilities. Perceptual disabilities, such as vision or neurological problems, i.e. blindness, low vision, or dyslexia are not considered in the bottom-line approach of big business. Big business does not consider small market segments; it is the essential role of libraries serving persons with disabilities to put forward the socially conscientious perspectives.

2.2 Synchronized Multimedia Integration Language (SMIL)

Multimedia is without a doubt the direction rich content will take. This offers persons with disabilities significant advantages, because the information can be represented through several sensory channels. Text, image, and audio synchronization can provide the user multiple channels to choose from to get the information. The blind user can select the audio channel, for example. However, the standards and the best practices for the production of content must address issues essential to persons with disabilities; the libraries are represented in the W3C working group charged with evolving this multimedia standard.

3 DAISY Standards

The Digital Accessible Information System (DAISY) Consortium has its roots in Libraries for the Blind Section of IFLA. The DAISY Consortium has gathered requirements from LBS patrons to identify the ideal reading systems for development. It has then integrated key experts in their employment to participate in W3C working groups, and in other technology development initiatives focused on information delivery. Once the W3C standards have been set, the DAISY Consortium applies those standards in the development of specifications specifically designed to serve persons with print disabilities.

The DAISY standards Based on W3C recommendations made their way into standards activities in the USA. The National Information Standards Organization (NISO) [4], with experts from the DAISY Consortium -- employees from LBS members evolved the initial DAISY work into the ANSI/NISO Z39.86-2005. The DAISY Consortium was named as the maintenance agency, the first time NISO has placed that responsibility on an international organization.

In addition, the DAISY Consortium is participating in the following standards developments:

- International Digital Publishing Forum (IDPF) [5]
- Consumer Electronics Association (CEA) [6]
- Organization for the Advancement of Structured Information Standards (OASIS) [7]
- International Telecommunications Union (ITU) [8]

4 Implementation of Standards

Unlike other standards organizations, the DAISY Consortium has a major focus on implementation of the standard. The Members of the DAISY Consortium are, for the most part, LBS members as well. The requirements are gathered from this group and from other related companies and organizations. With these requirements in hand, a long term "Road Map" is maintained to keep a clear vision of the implementation of the standards before the organizations. Standards, production tools, conforming content, and interoperable reading systems must all be in place for successful implementations.

It is instructive to detail some of the mechanisms that are currently in place:

- DAISY Knowledge Network: A Neuro Approach to Information
  A comprehensive community-based information system to share the collective knowledge of companies and organizations implementing the DAISY standards.
- DAISY OK: Interoperability between Reading Systems
  Producing the content once and having it render (play and present) on any conforming reading system is essential. DAISY OK self-certification makes this possible.
- Production Guidelines
  Knowing how to structure and produce content, convert books to a digital form, is one of the basic functions for the LBS organization. Building on the shared experiences of the organizations and putting them into guidelines are one of the services provided.
- Production Tools
  Identifying requirements in production tools and then setting out to make the tools, or encouraging vendors to produce tools that meet the requirements is fundamental in the overall process.

5 Services to Individuals

The Libraries for the Blind Section, in addition to contribution of expertise to the standards activities, must continuously engage in the provision of their basic services, which includes:

- Maintenance of the library catalog
  Develop and continuously update the accessible catalog of books and publications. In many cases, the catalog access is provided by library staff on the telephone.
- Delivery of Reading Materials
  Distribution to the patron through the post or internet of the titles and publications they have requested.
- Distribution of multiple formats, including powerful DAISY content
  Patrons require different formats, including braille, large print, and the DAISY format materials. The DAISY materials must be produced to meet the navigation and the high performance needs of the library patron.
- Descriptions of Graphical Content
  It is the libraries role to provide access to graphical content. This means that the important images, photos, flow charts, and diagrams must be described within the context of the publication in which it appears.
- Braille services
  Many of the libraries engage in the delivery of braille to their patrons. This is an invaluable service to those individuals who are braille readers.
- Tactile Graphics

Many graphical concepts are best communicated through tactile graphics. Having expertise in this area and selecting the right graphics to produce and deliver to the individual is essential.

- Mathematics, Music, and Specialized Content
  These types of items are extremely difficult to make accessible. They require special treatment. It is expected that the libraries will continue to produce this.

6 Development of a Global Library

Now that DAISY Standards have spread throughout the world, libraries are looking to share content across international borders. It is envisioned that production of a publication would happen once in the DAISY format and then share that title throughout the world. There are copyright laws and exceptions that must be addressed both nationally and on a global basis. In addition, the terms and conditions of collaboration among the libraries will need to be established.

7 Integration into the Mainstream

While the DAISY format is pervasive in the Libraries for the Blind Section, the mainstream of publishing has not embraced the standard. Competing standards and proprietary formats for digitally published materials abound; no format has yet risen to the forefront to be adopted. It is possible, and the DAISY Consortium is proposing, that the DAISY Standards could evolve to become the leading Standard for published content.

If a Standard for publishing, that is founded on accessibility could become the global standard for published content, then the function of libraries throughout the world would completely change. Of course, this is the vision we share.

8 Open Source, Raising the Bar, Breaking Down Barriers

The DAISY Consortium engages in extensive open source developments. This open source is in the areas of:

Conversion tools that

- Convert from inaccessible or marginally accessible formats to XML in the DAISY standard;
- Transform from one XML vocabulary to the DAISY XML;
- Upgrade utilities to move from previous versions of DAISY to the latest recommendation;
- Formatting utilities to output in formats such as braille and large print;

Validation utilities

Open source validation utilities that can be incorporated into other products or used as stand-alone products.

Reading systems
Open source reading systems, such as AMIS, that can be localized for any language.

Multimedia data model and SDK
It is expensive to build a comprehensive multimedia environment. The DAISY Consortium is breaking down the barriers to moving into software development by providing a comprehensive object oriented data model and a software SDK to be used in the building of products.

9 United Nations World Summit on the Information Society (WSIS)

The United Nations in 2003 and again in 2005 held the "World Summit on the Information Society" [9]. At both of these conferences, the DAISY Consortium and the libraries that make up the organization promoted the importance of including persons with disabilities in the design of our "Information Society". At both conferences, forums were held that brought together leaders in the disability community and raised a single voice that the United Nation's delegates could hear. This is yet another example of the essential role of libraries serving persons with disabilities in the Information Age; there is nobody else to raise the difficult issues.

10 Need for Continued Funding of Libraries Serving Persons with Disabilities

All the Full and Associate Members of the DAISY Consortium are non-profit organizations. Funding for the libraries in the world comes from a variety of sources. Please consider supporting your nation's library system that serves persons with disabilities. In addition, please consider supporting the DAISY Consortium through our online donation opportunities. Visit: `http://www.daisy.org` to learn more.

References

1. The DAISY Consortium: `http://www.daisy.org`
2. IFLA: `http://www.ifla.org/VII/s31/index.htm`
3. WAI `http://www.w3.org/WAI/`
4. NISO `http://www.niso.org/`
5. International Digital Publishing Forum (IDPF) `http://www.idpf.org/`
6. Consumer Electronics Association (CEA) `http://www.ce.org/`
7. Organization for the Advancement of Structured Information Standards (OASIS) `http://www.oasis-open.org`
8. International Telecommunications Union (ITU) `http://www.itu.int/home/`
9. World Summit on the Information Society `http://www.itu.int/wsis/`

Browsing Web Based Documents Through an Alternative Tree Interface: The WebTree Browser

Esmond Walshe and Barry McMullin

School of Electronic Engineering, Dublin City University, Glasnevin, Dublin 9, Ireland
esmond.walshe2@mail.dcu.ie, barry.mcmullin@dcu.ie
http://eaccess.rince.ie/

Abstract. The serial nature of speech technology seriously reduces the efficiency of blind individuals accessing web-based documents. Locating and manoeuvring to the required information can often be slow and laborious. This paper describes WebTree, a rather simple, yet highly customisable tree structured interface to web based documents, which provides page summaries based on the tree-like arrangement of the mark-up. The user dynamically controls how much of the document's tree hierarchy is to be exposed on a (virtual) screen at any given time. Thus, entire element sub-trees may be efficiently traversed with minimal difficulty. Methods for incorporating non-hierarchical elements (such as tables), are also discussed. In addition, an alternative search mechanism, which allows for the restriction of the search to specific mark-up elements, is examined. Finally, this paper includes the initial findings from user evaluation tests and provides some additional recommendations to increase the usability of the interface.

1 Introduction

The rapid growth of the World Wide Web has resulted in access to substantial quantities of written material previously inaccessible to those with little or no vision. Once accessibility guidelines, such as the Web content Accessibility Guidelines (WCAG)[1], are adhered to during the content design and authoring process, blind users can more easily navigate these websites and gain access to the required information through speech synthesis technology.

The WebTree system described in this paper does depend somewhat on an assumption that content complies with accessibility guidelines. Although levels of compliance are still quite low [1,2,3], there is good reason to expect an increase in website conformance. Web accessibility has been highlighted in legislation, such as section 508 of the Rehabilitation Act in the U.S.[2], and in case law.[3] The draft

1 http://www.w3.org/TR/1999/WAI-WEBCONTENT-19990505/

2 http://www.section508.gov/

3 *Reader's Guide to Sydney Olympics Accessibility Complaint:* http://www.contenu.nu/socog-PR.html

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 106–113, 2006.

code of practice[4] under the Irish Disability Act 2005[5] specifically requires public service websites to comply with level Double-A of the WCAG 1.0 guidelines.

There are huge differences in interaction methods involved when using purely audio output in comparison to visual interaction. A sighted user is generally capable of quickly scanning and selecting those items deemed as important page content. This is achieved through examining the spatial relationships between elements and through the use of visual cues, such as changes in font size, colour or emphasis. Unfortunately, due to the serial nature of speech technology, only a small fragment of the content (a brief segment of the audio stream) is directly perceivable at a time. Thus, to avoid reading the entire content to obtain a mental model of the page structure, additional in-page navigational approaches are potentially very helpful.

To obtain an overall picture of a web page, many current audio browsing solutions offer methods to navigate directly to the content enclosed within a subset of named mark-up elements, e.g., jump to the next/previous `<h*>` or `<table>` element. Sometimes a specialised summarisation function providing the reader with an overview of the content is also included. These range from views containing specific named elements, e.g., headers or hyperlinks[6,7], to document renderings based on the most common *word trigrams* in the text [4]. JAWS+MS-IE[8] also includes page summaries based on the first line/sentence within a paragraph, or entire paragraphs comprising specified words or phrases. Other summarisation solutions include Parente's *audio enriched links* mechanism [5] and the *Hearsay* system [6].

In addition, *tabular* material (a two dimensional grid of cells) poses serious issues for a serial interface. When reading horizontally across the row, the user may have to listen to the content of many cells before the required information is found. It may also be difficult to distinguish the contextual breaks between the cells. Raman proposed exploiting stereo (spatial audio) to indicate the location of the cell [7]. The first element of each row is spoken solely on the left channel; the rendering then progressively moves to the right, with the last element spoken solely on the right channel. Alternatively transformations are sometimes performed on the content to aid comprehension. The solution employed by JAWS+IE is to present each table cell in the sequence in which it appears in the mark-up with additional functionality to navigate from cell to cell. Yesilada et al also recommended a browser with such navigational functionality [8]. However, they also proposed performing *semantic* analysis to determine the context in which the cells appear, if presenting in a linear form.

---

[4] `http://tinyurl.com/o3rdp`

[5] `http://www.oireachtas.ie/viewdoc.asp?DocID=4338`

[6] JAWS for Windows:
`http://www.freedomscientific.com/fs_products/JAWS_HQ.asp`

[7] Home Page Reader (HPR):
`http://www-3.ibm.com/able/solution_offerings/hpr.html`

[8] Recent versions of JAWS have included specific features optimised for use with the Microsoft Internet Explorer web browser:
`http://hj.com/fs_products/software_jaws60fea.asp#Internet_Explorer`

2 Description of the WebTree System

WebTree has been developed to explore the use of an alternative, highly customisable tree structural approach to the auditory rendering of web based documents [9]. It provides alternative page summaries based on user-selected components of the underlying mark-up. The user dynamically controls how much of the document's tree hierarchy is exposed on a (virtual) screen at any given time. Thus, entire element sub-trees may be efficiently traversed with minimal difficulty. Through the customisation facility the user can decide which elements have their content automatically appearing in the display, or whether the element should be excluded altogether. This acts as a mechanism for generating alternative views of the same document. Similarly, styling can be associated with individual element types via an aural style sheet.

Although the current version of the system is optimised for XHTML[9] web pages, the approaches taken should be applicable to any tree structured mark-up document. XHTML was chosen due to the ease in parsing such files. However, the system stores the parsed document in a Document Object Model (DOM)[10] tree structure for internal manipulation. Both the navigation and display functions interface with the DOM structure to manipulate the document to the user's specifications. Therefore, once issues resulting from the accurate parsing of these documents are offset, the viewing of these pages should not be problematic[11].

The main document view is an abstraction derived from the hierarchical tree-like arrangement of mark-up elements. The view consists of a combination of buttons representing element sub trees, which when activated, expand or hide their content, and text content from elements already expanded. Each tree control element has two components. The first is a button that controls the expansion or hiding process. The second component provides the user with some contextual information about the element. This includes its name, e.g., `ul` and where possible the content of its `title` attribute. In the case of paragraphs, `<p>` elements, a certain number of characters contained in the element are exposed automatically. As an element is expanded, its child components appear beneath it in the virtual display.

Assigning tree expansion controls to *every* element would have reduced the overall usability of the system. Therefore, tree controls are reserved for *block* level XHTML elements. That is, *inline* elements are automatically expanded at the same time as the parent (block level) element. The display of tree-controls for a given element type can also be turned off to avoid higher level elements cluttering the display.

Tabular data poses additional difficulties for WebTree due to the two-dimensional (non-hierarchical) relationships between the cells. Once the `<table>` element is expanded, the individual elements contained within are presented in

[9] `http://www.w3.org/TR/xhtml-basic/`
[10] `http://www.w3.org/DOM/`
[11] However, JavaScript and "flash" components are not supported by the initial WebTree prototype.

the linear order in which they appear in the document. Lower level elements such as `<td>` or `<th>` elements are only exposed if their containing element (e.g., `<tr>`) is expanded. However, extra navigation functionality is provided to manoeuvre along the grid connections between the cells. As the user jumps to a new cell, the header information is not automatically announced. Instead, the user may jump directly to the header element, be it row or column to view its content. An *undo* command is available to reverse the jump. The user determines the amount of grid location information to be announced. This ranges from no contextual information, to speaking the beginning or end of the row or column, to announcing the grid coordinates for each cell.

Interaction with *forms* poses a number of difficulties when operating within a tree-based interface. Under WebTree, the presentation of form elements coincides with their linear organisation in the mark-up. All fields based on the `<input>` tag are automatically expanded as their parent element is expanded. Option lists (`<select>` elements), are treated in the same way as any ordinary tree component, except that the currently selected value of the list is included in the tree control information. Once expanded, the different options are presented as a list of checkboxes. The state of checkboxes and radio buttons may be altered by pressing the `enter` key. The user does not need to enter a specialised forms mode before access to the fields is granted under WebTree. Instead, as focus is placed on a form edit field, the application automatically alters the mapping of keystrokes from their prescribed WebTree settings back to the normal character insert function so that editing can take place. However, once focus has been removed from the form field/control, the key mappings revert to the WebTree specific keystroke mappings. A field is positioned on a separate line from any contained content to avoid confusion. As a form field is encountered during general reading, its presence is announced. When tabbing to the individual form fields, any associated `<label>` information is also read.

The WebTree application has also been augmented with specialised content search methods. These were included to experiment with a number of different navigational approaches. As well as searching for plain text, the user can search for a specific type of mark-up element, or limit the text search to specific element types. Thus if a document is properly marked up with structural elements, such as `<h*>` and `<em>`, navigation to the important areas of the page is facilitated. All element types are allowed as targets for the different search facilities.

3 Evaluation Methodology

To assess the usability of the WebTree application and the appropriateness of the tree approach to accessing web-based documents, five totally blind individuals in the age group of 21 to 31 were asked to evaluate the software. Four of the subjects had a good knowledge of computing in general, whereas the fifth subject had a more limited knowledge. All were reasonably proficient screen reader users. JAWS+IE was cited as being the predominant browsing platform used. Only two subjects reported having recently used any other audio browser application. Only one user had any prior experience using emacspeak[12], the environment under which WebTree operates. The levels of HTML knowledge varied significantly, ranging from one user whose knowledge was non-existent, to two users who had an advanced knowledge. All used the web on a regular basis for both work and leisure purposes.

Each user test was carried out separately, so that participants would not develop preconceived ideas of the system by listening to another test taking place in the background. Each participant was asked to fill out a preliminary questionnaire to determine some profile information, such as knowledge of HTML/XHTML elements and sophistication in using their current assistive technology. After filling out the preliminary questionnaire, the user was introduced to the WebTree application. They were initially afforded some time to familiarise themselves with the WebTree user interface. This was achieved by first reading the user manual (read through WebTree itself), and then through the completion of a short tutorial, which allowed them to experiment with the different types of functionality. There was no time limit placed on the familiarisation process, which typically lasted between one and two hours. Instead, the user was requested to indicate when they felt reasonably comfortable with the interface. At this point, participants were asked to perform a number of browsing tasks ranging from rather simple navigation tasks to more complex table navigation assignments. During the completion of these browsing tasks, the user had access to both the user manual and the tutorial as reference material. Participants were not encouraged to ask for help from the observer; instead, referring to the reference documents for help was recommended. However, if it was obvious that the user had become disorientated or confused, then assistance was offered.

Each participant was asked to complete seven browsing tasks in all. Four participants successfully completed all seven tasks; whilst the fifth subject's evaluation was curtailed due to unforeseen time constraints. After completing each task, the user filled out a short questionnaire to assess the usability of the application. The first four tasks involved navigating through a tree presentation of a snapshot mirror image of the eAccess website[13]. Users were asked to locate and navigate to a number of specified documents. The reason for including this number of similar tasks was to determine whether with practice the users' navigational efficiency improved. Also we wanted to ensure that if any usability issues occurred, they were related to the software and not usability problems with the website. The fifth task required subjects to search for different elements using both the specialised element search and the restricted incremental text search. To complete task 6, the user needed to locate and then fill out an XHTML form. Task 7 required the user to navigate through a number of table constructs using the special navigation functions. Once all the tasks were completed, each participant was asked to fill out two additional questionnaires, to establish usability data for the application and obtain user opinions about, and expectations of the types of functionality included.

[12] `http://emacspeak.sourceforge.net/`

[13] `http://eaccess.rince.ie/`

The structure of web pages tested differed greatly in element hierarchical complexity. This ranged from index pages containing multiple nested list elements (e.g., news pages and article indices), to large technical reports containing numerous data tables and sectional divisions.

4 Results

The lack of experience with WebTree prior to the evaluations meant it took some time before users became reasonably proficient with the system. A limited knowledge of both emacs and emacspeak also contributed to the problem. However, most users thought their proficiency would increase with a little practice. The user with a limited knowledge of HTML/XHTML mark-up elements initially had some trouble navigating documents. Once the mechanism of how such files are marked up was explained, he found the system quite usable. Another participant didn't like always having to expand elements to get to the required information. However, this problem was partly a symptom of the method he used to browse documents. Browsing for this user was achieved in a sequential manner expanding each element, as the information was sought. None of the search functionality was used by this subject to locate items. He also had issues with the time it took to navigate to the specified documents using this method. This problem could be partially attributed to a lack of knowledge concerning the layout of the website. As the users' knowledge of the website structure increased, the efficiency of their browsing methods was also enhanced. The other three users did not display any major difficulties in navigating with this system.

Most users stated that they liked the outline summary of the document that the tree structural view provided. This was especially the case when required to navigate through large documents. The fact that they could bypass complex items such as tables and just read the textual content they found appealing. In addition, four of the participants believed that displaying fragments of paragraph text when collapsed was beneficial in skimming the document. The fifth participant would rather that such a mechanism be optional. They all agreed that when `title`, or table `summary` attributes are provided, they should be included as part of the tree control text.

Although the presence of a certain element can be signalled using alternative speech properties assigned by an aural CSS, the effect of styling elements in this manner was not tested significantly. Possibly due to the learning curve required to become proficient with a completely new system, or the fact that all participants were used to reading with a single voice, users chose to have such cues turned off.

All users liked not having to specifically activate a specialised forms mode when entering form data. However, most users initially had some difficulty with the form construct. The problem was unrelated to the task of filling out information or locating individual fields. It was as a direct consequence of including structural elements inside form constructs to organise the content. Most users expected to find form input fields immediately after expansion, and were confused when a paragraph element was found instead. Therefore, we propose an additional function to expand an element's entire sub-tree to solve this problem.

Sometimes users became disorientated by being unsure which tree level was currently under focus. Although the user could establish an element's ancestor elements, an additional method such as a sound cue to signal a tree level change might be appropriate. Brewster et al experimented with using both hierarchical and combined earcons to portray hierarchical menus with some success [10]. However, further investigation is needed to determine whether this level of sound cue complexity is necessary in this case.

The only complaints in relation to the different types of search functionality were to do with the amount of audio feedback presented. In the case of large documents, searching may take a little time, thus notification that the search was still in operation was requested. Apart from that, users liked the different search technologies and thought them to be of benefit.

The table navigation functionality was well liked across the entire user group. The ability to jump to a table header element, peruse its content and return to the previous position was generally thought a good idea. However, an additional function to announce header content without leaving the current cell would be preferable. Users stated that they would rather not have header information read automatically as focus is placed on a cell. Instead, they would prefer it to be available on demand. In addition, a number of participants found the automatic announcing of row and column positional data quite irritating. A command to read this information at the users request would be more beneficial.

The emacs specific `customize` package was used to generate customisation views. However, none of the participants had any previous experience of working with emacs customisation buffers, and found them rather difficult to use. It was suggested that we scrap this method for providing customisation and implement an alternative system based on an XHTML form input system. Controlling the expansion of elements and the presence of tree controls, could be achieved by a set of checkboxes, or a multiple select list. Inserting items, such as line length, could be facilitated with text edit fields.

Although it may be unique to this small sample of users, it was evident that the regular browsing strategies of each user appeared to have an impact on how usable they perceived the system. Those users articulating that they usually navigated line-by-line or based on screen blocks such as paragraphs or by virtual pages, struggled at first to get to grips with the tree structural interface. However, those users that regularly use advanced navigation commands, such as navigating by element, or searching for required screen text, appeared to adapt to the system quite easily.

5 Conclusions

The initial user evaluations suggest that applying a tree approach to the viewing of web based documents is a viable solution. The greatest gains were observed in the case of large documents. However, it is expected that much greater usage with the application is needed before an accurate assessment of its usefulness is possible. As part of future work, we hope to examine the navigation of tables with cells spanning more than one row/column, under this system. Also, a method of unobtrusively notifying the user of their tree level position must be investigated.

Acknowledgments

The work described here received financial support provided from AIB PLC[14]. The work was carried out in the Research Institute for Networks and Communications Engineering (RINCE), established at DCU under the Programme for Research in Third Level Institutions operated by the Irish Higher Education Authority.

References

1. McMullin, B.: Users with disability need not apply? Web accessibility in Ireland. First Monday **7** (2002)
2. Marincu, C., McMullin, B.: A comparative assessment of web accessibility and technical standards conformance in four EU states. First Monday **9** (2004)
3. UK Cabinet Office: eAccessibility of public sector services in the European Union (2005)
4. Zajicek, M., Powell, C., Reeves, C.: Orientation of blind users on the world wide web. Contemporary Ergonomics (1998)
5. Parente, P.: Audio enriched links: web page previews for blind users. In: Assets '04: Proceedings of the 6th international ACM SIGACCESS conference on Computers and accessibility, New York, NY, USA, ACM Press (2004) 2–8
6. Ramakrishnan, I.V., Stent, A., Yang, G.: Hearsay: enabling audio browsing on hypertext content. In: Proceedings of the 13th international conference on the World Wide Web. (2004) 80–89
7. Raman, T.V.: Audio System For Technical Readings. PhD thesis, Cornell University, Computer Science Department (1994)
8. Yesilada, Y., Stevens, R., Goble, C.A., Hussein, S.: Rendering tables in audio: the interaction of structure and reading styles. In: The Sixth International ACM SIGCAPH Conference on Assistive Technologies (ASSETS). (2004) 16–23
9. Walshe, E., McMullin, B.: Accessing web based documents through a tree structural interface. In: Proceedings of ICCHP 2004. Springer Lecture Notes in Computer Science (LNCS), Springer (2004)
10. Brewster, S.A.: Using nonspeech sounds to provide navigation cues. ACM Transactions on Computer-Human Interaction (TOCHI) **5** (1998) 224–259

[14] `http://www.aib.ie/`

Web Pages for Blind People — Generating Web-Based Presentations by Means of Dialogue

Luděk Bártek and Ivan Kopeček

Faculty of Informatics, Masaryk University
Botanická 68a, 602 00 Brno
Czech Republic
{bartek, kopecek}@fi.muni.cz

Abstract. This paper deals with the problem of generating web pages by means of dialogue. The generated web presentations fully meet accessibility requirements. The basic approach is described and discussed. An illustrative example is presented as well.

1 Introduction

The Internet is now one of the most important sources of information for visually impaired and blind people. Their access to web information is supported by various assistive information technologies, like screen readers, special web browsers (see, e.g., [1,2]), standards of accessible web (see, e.g., [3,4,5]), etc.

In this situation, having one's own web presentation is advantageous and often a matter of prestige. Although there are many blind programmers and IT fans that are able to develop web presentations of high quality, for most blind users, creating their own web presentation is a not simple task and they usually need the help of a specialist.

In this paper, we describe and discuss the idea of developing web presentations through a dialogue, therefore in a way, that is fully feasible for a blind user. We concentrate on simple standard web presentations which can be easily generated. The side effect of the dialogue generation of the web pages is that the generated HTML code (see, e.g., [6]) fully meets the requirements of accessibility (see, e.g., [5,7]). The developed dialogue system for generating web-based presentations is implemented in VoiceXML (see, e.g., [8]).

2 Basic Concepts and Algorithmization

The dialogue getting the information about the generated web page can be divided into the following subdialogues:

- Subdialogue for describing the web page layout (see Section 2.1).
- Subdialogues for description of the web page graphics. The user can select the color of the background, position of the menu, the information whether the page will or will not have a heading, etc. (see Section 2.2).
- Subdialogues for description of the web page contents. The user can define the information that the page should present (text, images, sounds, etc.) (see Section 2.3).

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 114–119, 2006.

The entire web presentation can be generated by repeating the dialogue for every web page to be included.

2.1 Subdialogue Describing Web Page Layout

There are two most common layouts the user can select:

1. Simple information page – information and all links are included in the text (see Fig. 1.).

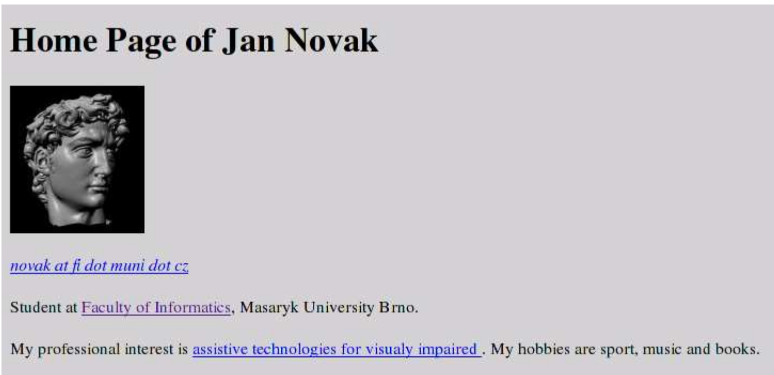


Fig. 1. Example of a simple information page

Fig. 2. Example of the information page with menu

2. Information page with menu. Most of the links in the page are located either on the side or the top of the page (see Fig. 2.).

The user can add the header and the footer of the page, select the height of the header and footer, select the width of the menu, etc. The dialogue strategy for this subdialogue asks the users to select the layout, then asks if they want to add a header and footer, and if the users want to use either a header or footer, asks them to enter the size of the particular parts of the generated page.

2.2 Subdialogue for Describing Web Page Graphics

The task of this subdialogue is to enable the user to select a web page background image or web page background color, the text color, etc.

The users are asked whether they want to set either the background color or put some image on the background. If the users choose a plain color, they are asked about its parameter (a predefined color, RGB specification, etc.) and if they selected image, they are asked for its URI (see, e.g., [9]).

2.3 Description of the Web Page Content

This subdialogue provides the description of the web page content. It enables the user to define the page information. This subdialogue will gather the following pieces of information:

1. Text information, typographic meta information (boundaries of the paragraphs, sections, ...).
2. Multimedia information (images, sounds, videos, etc.).
3. Links (both internal and external links are allowed).
4. Targets for internal links.
5. Meta-information – the author of the page, the title of the page, etc.

Whenever the non-textual information (image, video, etc.) is included, the user has to define the alternative description of the object to keep the web page blind-friendly (see, e.g., [10,11]). The collected information are inserted in a web page template, that is designed to fulfill the WCAG requirements.

2.4 The Web Page Generating

The generation can be performed in two ways. Either the server-side script can be used or the page can be generated using an embedded object, as described in the VoiceXML specification (see, e.g., [8]).

The advantage of using the embedded object is that the VoiceXML interpreter executes the embedded objects locally. The disadvantage is that the implementation of embedded objects is dependent on the VoiceXML interpreter.

The disadvantage of the server-side script is its dependency on the remote server. This can be reduced by running a web-server on a local machine, although this solution increases the system workload. The method of generating the web pages depends on the user preferences and can be easily modified in the systems configuration file.

An example of the dialogue generating the simple page shown in Fig. 1. follows:

System: Which web page layout do you want to use. Simple information page or page with menu?
User: Simple information page.
System: Do you want to modify the background of the page?
User: Yes.
System: What should be the background color?
User: Grey.
System: Should black text color be used as the default?
User: Yes.
System: Should some image be placed on the background?
User: No.
System: What is the first part of the page?
User: Heading, first level.
System: Enter the heading's text.
User: Home Page of Jan Novak.
System: What is the next part of the page?
User: Image.
System: What is the URI of the image?
User: www dot fi dot muni dot cz slash tilde novak slash img slash photo dot jpg
System: Enter the alternative description of the image.
User: Jan Novak's photo.
System: What is the next part of the page?
User: New paragraph.
System: What is the next part of the page?
User: Link.
System: Enter the link text.
User: novak at fi dot muni dot cz.
System: Enter the referring URI.
Users: mailto colon novak at fi dot muni dot cz
System: What is the next part of the page?
User: New paragraph.
System: What is the next part of the page?
User: Sentence.
System: Enter the sentence.
User: My professional interest is assistive technologies for visually impaired.
System: Does the sentence contain any link?
User: Yes.
System: Enter the text of the link.
User: assistive technologies for visually impaired
System: What is URI of the referred page?
User: www dot fi dot muni dot cz slash tilde bar slash hobbies dot html
System: What is the next part of the page?
User: Sentence.

System: Enter the sentence.
User: My hobbies are sport, music and books.
System: Does the sentence contain any link?
User: No.
System: What is the next part of the page?
User: End.
System: *Generates the requested page.*

3 Conclusions and Future Work

The idea generation of web-based presentations by means of dialogue has been discussed and tested by students and blind users during the 2005 course on Assitive technologies at Faculty of Informatics, Masaryk University Brno, and in cooperation with Teiresias Centre (Support Centre for Students with Special Needs) of Masaryk University.

Dialogue generation of web pages seems to be a promising technique that makes creating one's own web pages feasible even for blind Internet users, who are not overly interested in IT. This approach ensures that the generated pages fully meet the requirements of accessibility, which is another important advantage.

Our future work is aimed at testing and enhancing the efficiency of dialogue strategies and enlarging the set of basic standard profiles of web presentations. Another major task is to add a module which will provide the possible application of aesthetic criteria for the generated pages.

Acknowledgement

The authors are grateful to James Thomas for proofreading a draft of the paper and to the students and staff of the Support Centre for Students with Special Needs of Masaryk University for their advices, support and collaboration.

References

1. Webbie and the Accessible Programs, available at `http://www.webbie.uk`
2. Bártek, L., Kopeček, I.: Adapting Web-based Educational Systems for the Visually Impaired, In *Proceedings of the International Workshop on Combining Intelligent and Adaptive Hypermedia Methods/Techniques in Web-Based Education Systems*, Salzburg, 2005, pp. 11–16.
3. Authoring Tool Accessibility Guidelines 1.0, available at `http://www.w3.org/TR/ATAG10/`
4. User Agent Accessibility Guidelines 1.0, available at `http://www.w3.org/TR/UAAG10/`
5. HTML Techniques for Web Content Accessibility Guidelines 1.0, available at `http://www. w3.org/TR/WCAG10-HTML-TECHS/`
6. W3C HTML Home Page, available at `http://www.w3.org/MarkUp/`

7. Brewer, J.: Web accessibility hightlights and trends, In *ACM International Conference Proceeding Series; Vol 63, Proceedings of the 2004 international cross-disciplinary workshop on Web accessibility (W4A)*, New York City, New York, 2004, pp. 51-55.
8. Voice Extensible Markup Language (VoiceXML) 2.0, available at `http://www.w3.org/TR/ 2004/REC-voicexml20-20040316/`
9. Berners-Lee, T.: Universal Resource Identifiers in WWW, RFC 1630, June 1994.
10. Blind Friendly Web – Documentation on Accessible Web Design Guidelines for Severely Visually Impaired Users, available at `http://www.blindfriendly.cz/doc/bfw23en.zip`
11. Techniques for Web Content Accessibility Guidelines 1.0, available at `http://www.w3.org/TR/WCAG10-TECHS/`
12. Engelen, J., Evenepoel, F., Wesley. T.: The World WideWeb: A Real Revolution in Accessibility. In *Proceedings of ICCHP 1996*, Linz, Austria, July 1996, pp. 17-22.

Architecture for Personal Web Accessibility

Myriam Arrue, Markel Vigo, and Julio Abascal

University of the Basque Country
Informatika Fakultatea, Manuel Lardizabal 1, E-20018 Donostia, Spain
{myriam, markel, julio}@si.ehu.es

Abstract. Universal web accessibility is an inalienable objective to guarantee the civil right of all users to access to the Information Society and to avoid the digital gap. This is a long process based on the promulgation of inclusive laws, specification of accessibility guidelines, and development of adequate design methodologies and tools. To the short term, websites that do not fulfil the Universal Accessibility specifications can be used by specific groups of users that are not affected by the barriers present in these websites. Personal web accessibility concept focuses on the need of providing people with the adequate methods and tools to design websites so that adaptation and customization could be performed. This paper describes an architecture aiming to help any user to find, select and use websites that are currently accessible to her or him. In addition, it will aid web developers to create websites according to specific users' needs.

1 Introduction

The Universal Access concept is turning into something extremely significant for the emerging Information Society as it ensures access to the information in the World Wide Web by anyone, anywhere, and at any time [1]. Many efforts are being done to promote "universal access", "design for all", "inclusive design", "accessible design", etc., which are mostly overlapping terms referring to the same concept.

The use of on-line services includes currently a vast variety of services covering diverse needs. These services are accessed by means of different technologies, in diverse situations, and with different objectives. In addition, the fact that all citizens are potential users of on-line products has spread in the web community and developers are encouraged to design accessible sites. Consequently, the broad usage of websites is leading to significant changes in the traditional development process of software applications where Universal Access should be an essential quality target [2]. New websites development processes are starting to include Universal Accessibility as one of the main objectives.

However, many developers fail to design web applications according to Universal Access principles and they produce websites that can not be accessed by certain users, who may feel discriminated in such a technological society. One of the human groups that have large difficulties using on-line services is people with disabilities due to the fact that websites enclose usually many accessibility barriers.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 120 – 127, 2006.

A number of initiatives have been taken in order to overcome this situation including the promulgation, in some countries, of laws against electronic exclusion such as ADA, Section 508, etc. One of the most proactive initiatives is the Web Accessibility Initiative (WAI) [`http://www.w3.org/WAI/`] that was set up by the World Wide Web Consortium [`http://www.w3.org/`]. WAI elaborates and publishes well known guidelines and tools for evaluating web content accessibility.

In addition to being very useful for producing accessible web applications, these activities have extended the awareness of accessibility among web developers community. Even though they have not been sufficient, as in most cases good level of accessibility is not reached.

Universal accessibility is therefore an inalienable principle that must be achieved. Nevertheless, in practical terms, even if a specific website does not completely accomplish the universal accessibility goal, its accessibility level could be enough for some users that have physical, sensorial or cognitive restrictions. That means that personal accessibility is possible for a large number of users even when universal accessibility is not entirely achieved. However, in order to verify and profit from personal accessibility, each user should be able to find websites accessible for their own features. To this end, both users and web developers should be provided with methods and tools for ensuring personal web accessibility. Users should be able to find, select and use the websites accessible to her or him and web developers should be able to create websites according to user's specific characteristics. Therefore, methods and tools for evaluating and classifying websites based on their conformance to specific guidelines are necessary.

This paper presents an architecture which integrates the necessary methods, tools and techniques for achieving the personal web accessibility objective.

2 Architecture for Personal Web Accessibility

Personal web accessibility has a double objective. On the one hand, it aims to aid users searching for websites accessible for them. On the other hand, it aims to assist web developers in creating websites accessible for particular users' features. The first step towards the achievement of personal web accessibility is the development of flexible and efficient architectures in order to evaluate websites based on the users' personal web accessibility.

This type of architectures should integrate mechanisms for managing different users' needs and preferences by the use of Accessibility Profiles. In addition, they should be based on technologies which ensure that creating new accessibility guidelines based on personal experience, empirical data or assumptions can be easily performed. In this way, web accessibility evaluations based on specific characteristics of users would be possible.

2.1 Websites Evaluation According to Personal Accessibility

An Accessibility Profile is the set of guidelines associated to one or several user groups. Therefore, each Accessibility Profile determines the guidelines applicable for one user group. For instance, if we are taking into consideration the development of specific accessibility guidelines for users with different chromatic disabilities, it would be necessary to define the following ones:

- "do not rely on red colour" for people with protanopia disability.
- "do not rely on green colour" for people with deuteranopia disability.
- "do not rely on blue colour" for people with tritanopia disability.

It is necessary to provide methods and tools in order to define these specific guidelines for user groups and to make possible the integration of these types of accessibility profiles into the architecture of web accessibility evaluation tools.

First of all, it is necessary a flexible implementation of guidelines so that the statements in natural language and its implications on mark-up languages could be expressed in an easy way. This would ensure that new guidelines or updates of existing ones are correctly integrated into the architecture. Secondly, an interface to manage the accessibility guidelines and profiles is also necessary. This interface should offer the following options:

- Define new accessibility guidelines, based on empirical or personal criteria.
- Create and edit new accessibility profiles. Each accessibility profile will have associated the applicable accessibility guidelines.
- Create new accessibility profiles combining different existing profiles, i.e.: guidelines for elder users as they have sensorial, cognitive and physical disabilities in different ways.
- Publish developed accessibility profiles and guidelines so that other users could edit or extend them according to their needs.

An architecture containing the mentioned functionalities will make possible that web accessibility researchers could validate suggested new guidelines and could use them in order to create specific user groups focused websites.

2.2 Searching for Accessible Websites

Search engines are the most common way to find information when accessing the web. Consequently, the usage of this type of information retrieval systems is vertiginously increasing. However, these systems do not take web accessibility into account when producing their results, so end-users may find many accessibility barriers when trying to access the returned websites. Therefore, incorporating accessibility information in these information retrieval mechanisms would be very useful in order to overcome these situations as the searching task result items could be sorted by their accessibility level. Some related work has been proposed by Zhu et al. [3], in this case quality metrics were integrated into information retrieval systems.

Although web accessibility metrics are necessary for information retrieval processes, existing measures are not accurate enough. The most broadly accepted measures for web accessibility are the qualitative ones proposed in WCAG 1.0. In this document, three conformance levels, based on priorities of checkpoints, are proposed: A level which is obtained when all priority 1 checkpoints are satisfied in the website, AA level which is reached when all priority 1 and 2 checkpoints are satisfied and AAA level which is obtained when all priority 1, 2 and 3 checkpoints are fulfilled. In any other case, the website accessibility level would be 0. Therefore, a website fulfilling only all priority 1 checkpoints would obtain the same accessibility value than another website fulfilling all priority 1 checkpoints and almost all priority 2 checkpoints: both of them would get the A level conformance.

These criteria seem to be based in the assumption that if a webpage fails to accomplish one of the guidelines in a level, it is so un-accessible as if it fails to fulfil all of them. That is true for some users, but in general it is essential to have not only a reject/accept validation, but a more accurate graduation of the accessibility. Thus, as stated by Olsina and Rossi [4], defining quantitative metrics is necessary in order to overcome this situation. Moreover, they are essential in order to perform an adequate rating of websites.

This type of metrics has to be incorporated into the architecture so searching websites according to personal accessibility could be possible. In this way, the accessibility evaluation tools could obtain a quantitative accessibility value which could be useful for search engines in order to select and order websites based on users' personal accessibility.

2.3 Framework for Personal Web Accessibility

The Laboratory of HCI for Special Needs developed EvalAccess accessibility evaluation tool with the aim of automatically evaluate any set of web accessibility guidelines specified using a particular XML schema [5]. Therefore, the main characteristics of this tool are the following:

- A machine-understandable flexible accessibility guidelines implementation. This ensures an easy integration of new accessibility guidelines into the evaluation tool.
- An evaluation module able to automatically evaluate websites according to the defined accessibility guidelines. This module is independent of the accessibility guidelines to evaluate, which means that the guidelines are not built-in. In addition, the evaluation process can be customized, which means that it can be performed based on the guidelines associated to the user profiles.
- The resulting evaluation report is returned in a specific format XML Schema RXML (Report in XML). Due to the use of this XML language for the report it is possible to create machine readable and customized reports for users.
- A module for calculating personal accessibility quantitative metrics is integrated into the framework. This module gets all the necessary data for calculating the quantitative metrics from the resulting evaluation reports.
- An accessibility guidelines and profiles management module which provides the functionalities in order to define and modify personal accessibility profiles and guidelines.

The following figure, Figure 1, shows the architecture of the proposed framework.

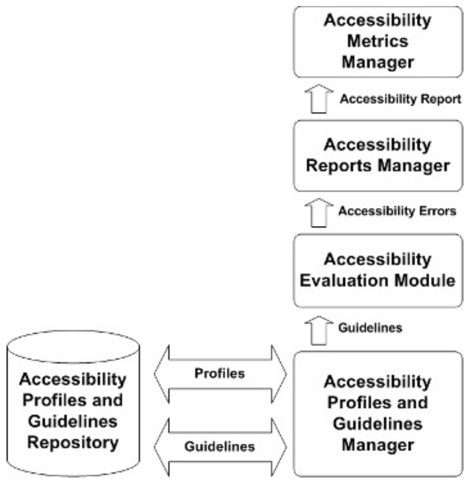


Fig. 1. Architecture of the described framework for Personal Web Accessibility

Accessibility Profiles and Guidelines Manager. The function of this module is to manage the different accessibility profiles and guidelines existing in the system. Depending on identification of the user, this module obtains his or her accessibility profile from the Accessibility Profiles and Guidelines Repository (APGR) in order to determine the accessibility guidelines associated to the user. Then, it processes this information for making it available to the Accessibility Evaluation Module.

The APGR is a native XML database which contains all the accessibility guidelines and profiles accordingly formatted for their easy management. The accessibility guidelines are formatted following a specific XML Schema [`http://www.w3.org/XML/Schema`]. This schema allows the specification of accessibility guidelines in a language based on XML, Guidelines in XML (GXML).

Due to this feature, it is possible to easily define new accessibility guidelines or update the existing ones. An application has been developed by the Laboratory of HCI for Special Needs in order to facilitate the definition of accessibility guidelines in GXML. This application can be locally executed by experts and the XML file generated is manually integrated into the database. Currently, an on-line and public version of this application is under development.

Accessibility Evaluation Module. This module evaluates the websites according to available accessibility guidelines which depends on the accessibility profile of the user. The detected accessibility errors are returned to the Accessibility Reports Manager.

Accessibility Reports Manager. The main objective of this module is to gather all the accessibility errors detected by the Accessibility Evaluation Module and produce a structured report. The resulting report is defined following a specifically produced XML Schema RXML (Report in XML). Due to the use of this XML language for the report it is possible to create machine readable and customized reports for users. A detailed description of RXML can be found in Abascal et al. [6].

Accessibility Metrics Manager. The Accessibility Metrics Manager module gets the accessibility errors report and applies the quantitative metrics to that information. Therefore, the results of quantitative metrics will be adapted to the user's capabilities as they are based on user-adapted accessibility evaluation results. For this purpose, some accessibility metrics have been defined and are described in detail in [7].

2.4 Framework Scenarios of Use

The implemented framework is useful to create guidelines based on personal experience, empirical data or assumptions. It can be useful in diverse scenarios of use. Some of these scenarios are the following:

- A research group wants to validate some proposed guidelines and designs a website according to them. Guidelines are coded in machine-understandable language and the website is automatically evaluated during the website development process as the framework contains an Accessibility Evaluation Module adaptable to all guidelines defined in different profiles.
- A specific user is aware of her or his capabilities when interacting with a website. She or he could have guidelines related to her or his requirements and could use them to search for most suitable websites according to her or his capabilities.
- The quantitative accessibility values obtained as a result of personal accessibility evaluation can be used by search engines or information retrieval processes in order to incorporate accessibility information in their results. This would be very useful as the searching tasks result items could be sorted depending on their accessibility level. A prototype that implements this scenario has been developed by the Laboratory of HCI for Special Needs [8]. This system has been developed taking advantage of Web Service technology, so it integrates two Web Services: EvalAccess [5] and Google [`http://www.google.com/apis/`].

Other scenarios not related to accessibility can be proposed. For instance:

- Personal accessibility mechanism can also be used to verify the presence or absence in a specific website of any characteristic specified as a guideline. That allows for instance the verification of the style book guidelines fulfilment in a corporate website.

3 Automatic Accessibility Evaluation Customization

The framework described in this paper can be considered to be adaptable to the end-user since it produces accessibility evaluation results based on users' specific characteristics. Although generally accessible (fulfilling general purpose guidelines) web applications can be developed in order to easily perform these tasks, their interface can not be completely accessible for some user groups. Then, they could hardly specify any specific guidelines or profiles. Moreover, defining the set of accessibility guidelines which should be associated with their accessibility profile is a difficult task for any user. This task is currently performed by experts. Nevertheless, automatic customization of user specific accessibility will be a good option in this context, leading to user accessibility adaptation based on personal capabilities.

Universal Access concept requires that software systems could tailor diverse set of users. This has stimulated interest in user interfaces able to adapt to users. User profiles have been extensively used for achieving this purpose as they contain user information which could be useful for adaptation processes [9].

The framework described in this paper could be extended by adding the functionality of managing personal user profiles. In this sense, the system could infer the accessibility profile of one specific user from the information contained in his/her user profile.

The Composite Capabilities/Preferences Profiles (CC/PP) provide a framework [`http://www.w3.org/Mobile/CCPP/`], based on XML/RDF, for defining user and device profiles for adaptation of web systems in terms of their capabilities and preferences. In addition, its vocabulary can be easily extended in order to accommodate new features.

Even if some of the parameters used to characterize a user in CC/PP can be automatically extracted in the interaction process (such as the preferred browser or the operating system), the most meaningful data for interaction, such as the type of physical, sensorial or cognitive restrictions cannot. Therefore, a separate phase, usually previous to the interaction, serves to ask the user about key interaction data. This information is used to characterize the user profile. Traditional user modelling and adaptation can be applied to provide interfaces adapted to each user needs. That includes defining the parameters that are relevant and observable, and the range of values that they can take. This definition is connected with the capability of adaptation of the interface. It is no sense to define parameters or ranges of values that cannot be considered for adaptation. It is also convenient to define stereotypes of users that have in common a subset of characteristics. That allows to generalize some adaptations and to provide a first interface to people not having a complete profile. On the other hand, many interaction parameters may change with time. It is necessary to be able to dynamically adapt to these changes. In addition to rules that assign specific interface behaviour to each profile, mechanisms to reason about contradicting data, stereotypes assignation, etc., are also needed.

Therefore, considering the integration of user profiles in the existing framework for accessibility evaluation is an important step forward personal accessibility. This framework seems to be useful in order to achieve the aimed automatic customization of web accessibility evaluation.

4 Conclusions

Universal accessibility is an inalienable objective that must be reached in order to guarantee the civil right of every one to access the net and to avoid the digital gap. In addition, individuals should be provided with tools that allow them to find the existing websites that fulfill their own accessibility needs, that is, to ensure Personal Accessibility.

In this paper we have discussed our approach to Personal Accessibility, including the overall architecture and the modules currently developed. In other words, the proposed framework is an extensible architecture for web customization. An interface is provided so that personal guidelines defined in GXML could be designed, edited and stored by any user, no matter the knowledge of XML one has. Once guidelines are specified the guideline-independent tool performs evaluations according to them and errors are reported. These machine understandable reports make easy obtaining quantitative accessibility metrics which are necessary for information retrieval processes in order to classify websites according to their accessibility.

Finally, the forthcoming user adaptability module is described which will be easily integrated, due to the flexibility of the framework.

References

1. Stephanidis, C. and Savidis, A. (2001). Universal Access in the Information Society: Methods, Tools, and Interaction Technologies. International Journal of Universal Access in Information Society (UAIS). Vol. 1, No. 1, pp. 40-55.
2. Savidis, A. and Stephanidis, C. (2004). Unified user interface development: the software engineering of universally accessible interactions. International Journal of Universal Access in Information Society (UAIS). Vol. 3, No. 3-4, pp. 165-193.
3. Zhu, X. and Gauch, S. (2000). Incorporating Quality Metrics in Centralize/Distributed Information Retrieval on the World Wide Web. Proceedings of the 23rd International ACM SIGIR Conference, SIGIR '00, pp. 288-295.
4. Olsina, L. and Rossi, G. (2002). Measuring Web Application Quality with WebQEM. IEEE Multimedia. Vol. 9, No. 4, 20-29.
5. Abascal, J., Arrue, M., Fajardo, I., Garay, N. and Tomas, J. (2004). Use of Guidelines to automatically verify web accessibility. International Journal on Universal Access in the Information Society (UAIS), Vol. 3, No. 1, pp. 71-79.
6. Abascal, J., Arrue, M., Garay, N. and Tomas, J. (2003). A Web Service for Automatic Accessibility Analysis of Web Pages Based on the Use of XML Structures. Proceedings of HCI'03. pp. 925-929.Lawrence Erlbaum Associates.
7. Arrue, M., Vigo, M. and Abascal, J. (2005). Quantitative Metrics for Web Accessibility Evaluation. Proceedings of the ICWE 2005 Workshop on Web Metrics and Measurement. University of Wollongong School of IT and Computer Science.
8. Arrue, M., Vigo, M. and Abascal, J. (2006). Towards Web Accessibility Awareness in Information Retrieval systems. Technical Report EHU-KAT-IK-05-06, The University of the Basque Country.
9. Koutsabatis, P., Darzentas, J. S., Konstantopoulus, M., Spyrou, T. and Darzentas, J. (2001). Specifications for translating user models to communications channels. IRIS Consortium, Deliverable D0301.

Dynamically Generated Scalable Vector Graphics (SVG) for Barrier-Free Web-Applications

Kerstin Altmanninger and Wolfram Wöß

Institute for Applied Knowledge Processing (FAW)
Johannes Kepler University Linz, Austria
kerstin.altmanninger@students.jku.at,
wolfram.woess@jku.at

Abstract. Many graphics used in Web pages are very attractive to the eye and useful for many people. However, with the extensive use of pixel graphics such as charts without textual description or image maps, Web pages are encountering an ever increasing amount of barriers. There are many developments to aid people with visual impairments to gain access to graphics on the Web but most of these techniques are not universally applicable to other disabilities. It is essential that future development concentrates on accommodating all kinds of disabilities. The use of Scalable Vector Graphics (SVG) provides new possibilities as well as new challenges for the accessibility of Web sites. Consequently, this paper introduces a solution to make all graphics accessible to each user-group, and visualizes them in the resultant prototype *Access2Graphics*.

Keywords: Scalable Vector Graphics (SVG), accessibility, elderly/disabled people, barrier-free graphics.

1 Introduction

Surfing the Internet is not always as simple as a single click away. Disabled people often encounter many barriers which can be overwhelming for them. To make the World Wide Web more accessible to the broader community the Web Accessibility Initiative (WAI) works with organizations around the world to develop strategies, guidelines and resources [1]. Barrier-free Web design offers handicapped users the ability to perceive, understand, navigate and interact in the WWW environment with as few accessibility issues as possible. To achieve this goal, the World Wide Web Consortium (W3C) has specified a set of guidelines that form a foundation on which further developments should follow [2]. Accessibility does not solely concentrate on making the Web *more accessible* for disabled people. It also incorporates availability of Web content and enhancing the Web experience for as many people as possible, regardless if their disabilities is related to vision, motor, learning impairment or if someone has no disability at all. Availability problems of Web content can be caused by slow network connections, slow computer processing or software agents that lack understanding of how a human views and interacts with a Web page. Mobile phones and PDAs with limited bandwidth, memory and screen resolution also entail limitations.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 128 – 135, 2006.

Created to be an information portal, the Web should be made accessible for as many people as possible. That does not mean that Web pages need to be designed in a non-attractive way, lacking colors and pictures. Working with style sheets, dynamically generated information and other common Web design techniques leads to a more attractive Web and allows at the same time the consideration of every person's specific requirements. For instance, cognitive handicapped and elderly people find that sites employing a lot of graphics are much easier to use and to understand because they support them to focus in quickly on their intended action. The fact that blind people and those with low vision cannot see graphics at all or not so clearly should not be a motivation to avoid graphics in the Web. In contrast to raster graphics such as PNG or JPEG which many disabled people find difficult to work with, Scalable Vector Graphics (SVG) are very beneficial. Providing an alternative text for graphics as it is proposed in the Web Content Accessibility Guidelines [2] is not always satisfactory for people with visual impairments. Physically disabled users often have problems with image maps visualized with raster graphics because they are not scalable and therefore it may be difficult to select the right field. These user groups demand a textual based graphic format. The vector graphics format SVG is predestined to tackle these shortcomings. Using SVG offers a number of features to make graphics on the Web more accessible than this is currently possible with pixel graphics. Describing those features the W3C Note Accessibility Features of SVG [3] present many techniques to benefit users with low vision, color blindness, blindness or users of assistive technologies.

As previously mentioned some related work has been done in this field by developing World Wide Web Consortium (W3C) guidelines [2,3]. Recent publications concerning access to SVG for visually impaired people focus on specialized areas, specifically extracting meta information [4] and visualization of tactile maps [5,6,7]. Research on exploring Scalable Vector Graphics for visually impaired users in general is becoming more common [8]. Thus, far studies mostly concentrate on blind, color blind or people with low vision presenting a static graphic solution and transforming it for the appropriate output device. This paper is going a step further by considering every kind of disability and its relationship to graphics. It provides an application for the dynamic generation of a Scalable Vector Graphics specific to each user type with none, one or more disabilities.

The paper is structured as follows: Section 2 gives a brief introduction of the W3C standard SVG. In Section 3 user groups with characteristic disabilities and their derived requirements concerning graphics are discussed. Consequently, in Section 4 the user adapted dynamic image generation process will be described and the prototype of the *Access2Graphics* application is presented. Finally, Section 5 concludes the paper and gives an outlook on further research activities.

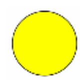
2 Scalable Vector Graphics (SVG)

SVG, an XML-based two-dimensional graphics standard recommended by the W3C, can aid in making graphics accessible for everybody. At its core, SVG is plain text with the inclusion of vectors that retain their resolution at any zoom. Furthermore, SVG allows Web developers and designers to implement dynamically generated, high-quality graphics from real-time data with precise structural and visual control. With this powerful new technology, SVG developers have the possibility to create a new generation of Web applications based on data-driven, interactive and personalized graphics [9], currently only possible after installing the Adobe SVG plug-in.

For graphics specification SVG offers a set of basic shape elements such as lines, circles, ellipses, polygons, poly-lines and rectangles. Additionally, it is possible to include external shapes and text in the graphic.

The sample source code below demonstrates some core features of SVG. Firstly, the standard provides a special <g>-tag to group elements. This feature enables a Web designer to arrange graphics in layers, giving the user the opportunity to select favored layers and to view specifically requested information. Secondly, SVG provides an optional title and description tag for each graphical layer which can be used to store meta information.

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE svg PUBLIC "-//W3C//DTD SVG 1.1//EN"
  "http://www.w3.org/Graphics/SVG/1.1/DTD/svg11.dtd">
<svg width="100%" height="100%">
  <title>Smiley</title>
  <desc>The picture shows a yellow smiling face.</desc>
  <g>
    <title>Face</title>
    <desc>The face is represented by a yellow circle.</desc>
    <circle cx="60" cy="60" r="40"
     style="fill:yellow;stroke:black;stroke-width:1;"/>
  </g>
  <g>
    <title>Eyes, nose and mouth</title>
    <desc>The facial expression is displayed in black color.</desc>
    <g style="fill:black;stroke:black;stroke-width:1">
      <title>Eyes</title>
      <circle cx="48" cy="44" r="3"/>
      <circle cx="72" cy="44" r="3"/>
    </g>
    <g>
      <title>Nose</title>
      <line x1="60" y1="50" x2="60" y2="70"
       style="stroke:black;stroke-with:2;"/>
    </g>
    <g>
      <title>Mouth</title>
      <ellipse cx="60" cy="75" rx="15" ry="7"
       style="fill:black;stroke:black;"/>
      <ellipse cx="60" cy="71" rx="15" ry="7"
      style="fill:yellow;"/>
    </g>
  </g>
</svg>
```



The fact that SVG is an XML-based language has many advantages. It is useful in combination with scripting and style sheet languages and in addition, SVG files can be generated with XSL Transformations.

The source code above describes a simple graphic – a ‘smiley’ – composed of circles, ellipses and a line. However, SVG can be used in much more complex applications for generating graphs and charts resulting from database queries [10], for designing geographical maps, image maps, Web sites, interactive and/or animated graphics as well as for images.

In the following chapter it will be explained, how the features of SVG can be used to make graphics accessible for different user groups.

3 Requirements of Specific User-Groups

People with visual disabilities, hearing impairments, physical disabilities, speech impediments or cognitive disabilities are in many cases confronted with insuperable barriers when they are working with Web applications. Assistive devices like alternative keyboards or switches, Braille displays and screen reading software are supporting tools to reach a higher level of access to Web content, but there are still unsolved problems concerning graphics accessibility. How people with different kinds of disabilities are using the Web in general is discussed in many studies [11]. When focusing on the barriers of common raster graphics, the following user groups can be identified:

- people with visual disabilities (blindness, low vision, color deficiencies)
- people with hearing impairments
- people with physical disabilities
- people with cognitive disabilities

Blindness. Blind people need an alternative text for images. The Web Content Accessibility Guidelines postulate in guideline one that graphics should have an alternative text [2]. However, for blind people this description is often not sufficient. Normally graphics are composed of many different layers and therefore SVG provides text equivalents on each of these layers.

In generated SVG graphs or charts it is preferable to present information in a table form only. Some approaches made it possible to gain access to such graphics [5] and geographical maps [7] via a tactile display. This output device can be used in combination with textual equivalents that are accessible for blind people through voice output and Braille text. In all, the amount and meaning of meta information included in a graphics determines its degree of accessibility.

Low vision. Poor acuity (vision that is not sharp), tunnel vision (seeing only the middle of the visual field), central field loss (seeing only the edges of the visual field) or clouded vision describes the forms of low vision. To achieve access to graphics people with low vision use screen magnifiers, screen enhancement software and screen readers. In this area user defined style sheets are growing more and more important. SVG is appropriate for high level scaling but as a drawback for people with low vision it is very difficult to gain an overview of an image that has been excessively scaled. Therefore also for users with low vision it is vital that a graphics includes as much useful textual meta information as possible.

Color deficiencies. One of the most common visual disabilities is color blindness, a lack of perceptual sensitivity to certain colors. There exist two widespread forms of color blindness, red-green and yellow-blue. To avoid problems with the definition of color graphics a textual description of the graphics or a user defined style sheet provides assistance.

Hearing impairments. SVG allows graphics to include voice output. For people with impaired hearing acoustical information must be provided alternatively, e.g. as text.

Physical disabilities. Motor or physical disabilities can affect someone using a keyboard or mouse. Static graphics have no real effect on information for this user group but if a graphics is interactive, problems can result. Therefore it is favorable for interactive image maps to have sufficient space between the different links and that the icons themselves are larger in size.

Cognitive disabilities. Difficulties with reading, understanding abstract concepts, and performing mathematical calculations are examples for cognitive disabilities. The utilization of graphics is very important for cognitive disabled people as it can significantly improve the understanding of a Web site's content. SVG may also incorporate text that needs to be as clear as possible with the simplest language appropriate. For example, charts should not contain abbreviations without additional explanation.

Vision loss, hand tremor, short-term memory loss and variations of cognitive disabilities are often accompanied with aging. Such related conditions are frequently a combination of many of the aforementioned disabilities and so will not be considered as a specific type of disability.

4 User Adapted Dynamic Image Generation

Considering the different kinds of disabilities mentioned above, it is a fact that one and the same graphics cannot satisfy individual user requirements. Therefore in this paper the application *Access2Graphics* is introduced, which is based on user profiles and the graphics standard SVG. *Aceess2Graphics* aims at enabling a dynamic generation of individually adapted and consequently accessible graphics applicable in any WWW environment. In the following sections the generation and adaptation process as well as the implementation prototype are presented.

4.1 Access2Graphics Application

Figure 1 illustrates the workflow and generation process for context and user dependent graphics. The starting point is the user characterized by a profile which specifies the user's disabilities. By default the application assumes that a user has no disability and no special demands concerning the output device. If the user requests a Web page containing graphics, in the next step a corresponding database query is generated. At this stage the result set of the query, which is already in SVG format, considers the disability specification of the user profile. In a further step, *Access2Graphics* assembles the Web page containing the user adapted images by considering both, the additional disability restrictions and the user's output device, ie. monitor, PDA or mobile phone. Additional disability restrictions are for example color blindness, making it necessary to switch an image from color to black and white. *Access2Graphics* uses filters for this task. For visually impaired people, output can be processed by a Braille device. Tactile displays are not supported by the actual version of the prototype. Additionally, there is the option to obtain voice output as well.

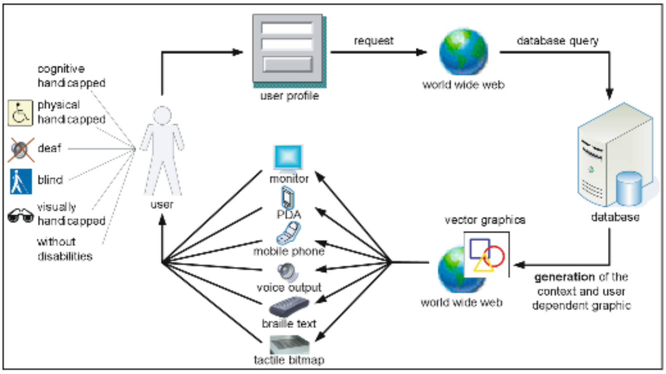


Fig. 1. Dynamic generation of context and user dependent graphics

One of the key advantages of this process is that the SVG information required for the different kinds of users is stored in the database. Therefore the basic SVG information is generated once and can then be adapted corresponding to specific demands. This is in contrast to existing solutions where the SVG information normally considers only one user demand. Of course the flexibility of *Access2Graphics* directly depends on the SVG data quality stored in the database.

4.2 Implementation Prototype

The *Access2Graphics* prototype was implemented with the server-side scripting language PHP in combination with a MySQL database and an Apache Web server. The following aspects are responsible for this decision. The MySQL database is open source software in contrast to the common databases Oracle and Microsoft SQL Server. It also has many advantages concerning performance, usability, support of larger databases and compatibility with common operating systems (Windows, Linux, Mac OS and most Unix derivatives). The scripting language PHP is designed to support databases, is secure, fast and support is widely available. Moreover, the communication between MySQL and PHP works very well [12].

As mentioned above, many types of SVG graphics can be generated. This prototype concentrates on automatic generation of charts, image maps and any kind of images. The *Access2Graphics* tool does not claim to implement each graphics type creatable with SVG but aims on a demonstration of what is possible.

To get started a user must define his/her user profile as shown in Figure 2. Accordingly, he/she may select more then one checkbox of the user specific settings. Additional adjustments in output devices such as voice output and grayscale graphic display can be selected under further settings. The user is also allowed to specify an appropriate output device different from the default setting "monitor". After completing the user profile form SVG images included in Web pages will be generated via a database query. To make this generation process possible meta information about the mapping and transformation of different graphics types to corresponding user profile specifications is stored in the database. Of course, graphics data itself is part of the database, e.g. tables to generate charts or source code for images. SVG provides the ability to fragment a graphics into several layers. These layers are also called groups, which have to be declared with meta information containing a textual description of the layers content. Consequently an appropriate number of layers improve the accessibility because they cover a structured and extensive description of the image. In the above code example the tags `<title>` and `<desc>` explain that a smiley is a face with eyes, nose and a smiling mouth.

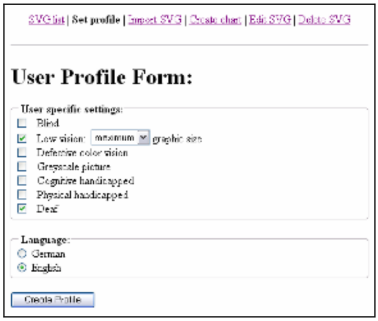


Fig. 2. Access2Graphics user profile form

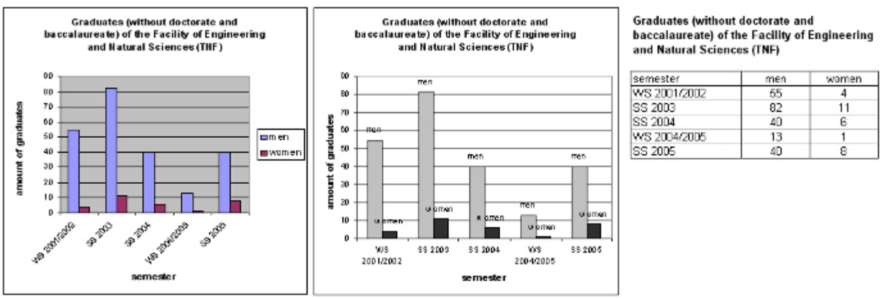


Fig. 3. Examples for accessible SVG graphics

Figure 3 shows three examples for accessible SVG graphics. The left bar chart is the “original” one which is generated for a user without any user profile specifications. The center chart is generated for a person with defective color vision. These users have the option to design a user-dependent color graphic or simply display the graphic as a grayscale image. For a blind user group a SVG bar chart would be senseless. Hence, content of the database will be provided in a HTML-table.

5 Conclusion and Further Work

Visual, cognitive, physical and hearing impairments should not longer cause a barrier with respect to graphics that they currently do. It is a fact that pixel graphics provided without an alternative text are useless for blind people and bar charts that seem to have similar colors for different bars are difficult to identify by color blind people. To overcome these problems *Access2Graphics* makes images dynamically accessible for disabled Web users, thus, offering a new way of exploring graphics for them by adding together pieces of former models. A further advantage of the presented approach is that it does not concentrate solely on visually impaired people as this is the case in preceding studies. The introduced prototype *Access2Graphics* visualizes some basic examples using the W3C standard SVG.

Further work in this area is intended and will concentrate on an advanced SVG parsing process. Consequently, graphics individualization should be refined in order to accurately meet the specific user requirements. The parsing process will also be extended with additional graphic types. To sum up, in the future the application *Access2Graphics* should have the ability to import, parse, manage and display as many different SVG images as possible.

References

1. WAI – Web Accessibility Initiative: http://www.w3.org/WAI (2005)
2. W3C – World Wide Web Consortium: Web Content Accessibility Guidelines 1.0, W3C Recommendation 5-May-1999, http://www.w3.org/TR/WAI-WEBCONTENT (1999)
3. W3C – World Wide Web Consortium: Accessibility Features of SVG, W3C Note 7 August 2000, http://www.w3.org/TR/2000/NOTE-SVG-access-20000807 (2000)
4. Herman, I., Dardailler, D.: SVG Linearization and Accessibilitiy, http://www.svgopen.org/2002/papers/herman_dardailler__svg_linearization_and_accessibility (2002)
5. Rotard, M., Ertl, T.: Tactile Access to Scalable Vector Graphics for People with Visual Impairment, www.svgopen.org/2004/papers/TactileAccessToSVG (2004)
6. Bulatov, V., Gardner, J.A.: Making Graphics Accessible, http://www.svgopen.org/2004/papers/SVGOpen2004MakingGraphicsAccessible (2004)
7. Campin, B., McCurdy, W., Brunet, L., Siekierska, E.: SVG Maps for People with Visual Impairment, Proceedings of SVG Open, Vancouver, Canada (2003)
8. Rotard, M., Otte, K., Ertl, T.: Exploring Scalable Vector Graphics for Visually Impaired Users. In: Klaus, J., Miesenberger, K., Burger, D., Zagler, W. (eds.): Computers Helping People with Special Needs. Lecture Notes in Computer Science, Vol. 3118. Springer (2004)
9. Adobe Systems Incorporated: Scalable Vector Graphics, http://www.adobe.com/svg (2006)
10. González, G., Dalal, G.: Generating SVG Graphs and Charts from Database Queries http://www.svgopen.org/2003/papers/GraphsAndChartsFromQueries/index.html (2003)
11. W3C – World Wide Web Consortium: How People with Disabilities Use the Web, W3C Working Draft, http://www.w3.org/WAI/EO/Drafts/PWD-Use-Web/20041210 (2004)
12. Watt, A., Lilley, C., Ayers, D.J., George, R., Wenz, C., Hauser, T., Lindsey, K., Gustavsson, N.: SVG Unleashed, Sams Publishing (2003)

Designing a Web Page Considering the Interaction Characteristics of the Hard-of-Hearing

Miki Namatame[1], Tomoyuki Nishioka[1], and Muneo Kitajima[2]

[1] Tsukuba University of Technology,
4-3-15 Amakubo Tsukuba Ibaraki 305-8520 Japan
{miki, nishioka}@a.tsukuba-tech.ac.jp
http://www.tsukuba-tech.ac.jp/

[2] National Institute of Advanced Industrial Science and Technology (AIST),
1-1-1 Higashi Tsukuba Ibaraki 305-8566 Japan
kitajima@ni.aist.go.jp
http://staff.aist.go.jp/kitajima.muneo/

Abstract. The purpose of this paper is to report a case study involving the successful redesigning of a Web page that was problematic for the hard-of-hearing. We found in our previous eye-tracking studies [1,2,3] that the page in question presented serious usability problems for the hard-of-hearing. Namely, the performance of hard-of-hearing participants was inferior to that of the hearing in terms of the following four performance measures: 1) scan patterns, 2) the number of errors and the time necessary to select the correct link, 3) the amount of time necessary to select a link, and 4) the types of selected links. We conjectured that these differences occurred because the informational organization of the original Web page was difficult for the hard-of-hearing to understand. Considering the Web interaction characteristics of the hard-of-hearing, we redesigned the page in two ways: 1) adding vertical lines that should function as visual support enabling the hard-of-hearing to grasp the informational structure easily, and 2) replacing difficult-to-understand labels with comprehensible representations. Observation of eye movements for the redesigned page revealed that the abovementioned differences disappeared, indicating that the redesign was successful. We believe that this case study exemplifies the successful redesigning of Web pages to make them more accessible to hard-of-hearing users.

1 Introduction

Recently, with continued advances in information technology, an ever-growing amount of information has accumulated on the World Wide Web. At the same time, the need to make the information accessible to any person who needs it has become a serious issue. This paper focuses on Web contents accessibility for the hard-of-hearing. This project was motivated by the fact that the first two authors, who are educators of hard-of-hearing persons, perceived in daily

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 136–143, 2006.

classes that hard-of-hearing students interact with Web pages differently than hearing students do. These differences in material usage suggest that the hard-of-hearing may not effectively use educational materials that are not designed appropriately from their viewpoint. Therefore, as educators, the authors need to create Web-based educational materials that are accessible to hard-of-hearing users.

Consideration of Web content accessibility for the hard-of-hearing is usually limited to the issue of translating auditory information into sign language and/or text annotation (i.e. translating physically inaccessible representation of Web contents so that it is physically accessible). However, studies reveal that this simple translation is not sufficient to allow the hard-of-hearing to attain *true* accessibility. First, although a number of efforts have sought to utilize sign language in the Web environment (e.g. [4]), they are not always effective, at least in Japan, since only 10% to 20% of the hard-of-hearing can use the mother-tongue sign language. Second, the use of text annotation alone is not sufficient for the hard-of-hearing to understand appropriately the meaning of Web contents.

Using Web-based interactive materials seems effective for education since they allow the creator to control the presentation of the content. An excellent example is the Web-based educational material for film production developed by [5], which provides sign language for various countries. However, it is necessary to accumulate more knowledge about how the hard-of-hearing use the Web in a broader context, not restricted to the use of sign language, in order to help them gain the full benefits of the Web environment.

This paper is organized as follows. Section 2 describes our previous studies [1,2,3] that presented Web interaction characteristics of the hard-of-hearing. Section 3 reports on a case study concerning the redesign of a problematic Web page for the hard-of-hearing by considering the Web interaction characteristics of the hard-of-hearing. The redesign was proven successful in terms of various performance measures. We believe our practice successfully creates effective Web-based interactive materials for the hard-of-hearing.

2 Web Interaction Characteristics of Hard-of-Hearing

This section reviews our previous studies [1,2,3] that revealed differences between the Web-browsing behavior of hard-of-hearing persons and that of hearing persons when they accomplished a task on an experimental Web page that simulated a then-existing automobile site. The participants were asked to locate a page that described a designated car model, and to choose a favorite color for it. We recorded their link selections and eye movements, and analyzed the data from the top page.

2.1 Task

The task was to locate a page that described car model Z4, and to choose a favorite color for it. The subjects were given the following instruction: “Please choose your favorite color for the car model Z4." The task was performed on an experimental Web site modified from an actual automobile Web site. The left portion of Figure 5 illustrates the top page, which consisted of five columns (four content columns and one news column) and a field at the bottom of the page where the names of car models were listed. Content columns had a heading at the top, a picture with promotional text in the middle, and a list of topics at the bottom. An important feature of this page was that the contents were organized vertically. Successful task performance required correct understanding of the page layout, since the column boundaries were not clearly defined.

2.2 Differences in Performance

We examined four performance measures to understand Web interaction characteristics of the hard-of-hearing: 1) scan patterns, 2) the number of errors and the time necessary to select the correct link, 3) the amount of time necessary to select a link, and 4) the types of links the participants selected. The following subsections briefly discuss how the hard-of-hearing's performance differed from that of the hearing in terms of these performance measures.

Scan Patterns. Figure 1 compares the scan path of one hard-of-hearing participant (left) with that of one hearing participant (right). The two participants took approximately the same amount of time to accomplish the task. However, it is clear from the figures that their scan paths were completely different. The left scan path (hard-of-hearing) does not reveal any clear pattern of scanning. On the contrary, the right scan path (hearing) indicates a vertically aligned scan path, consistent with the underlying semantic structure of the page.

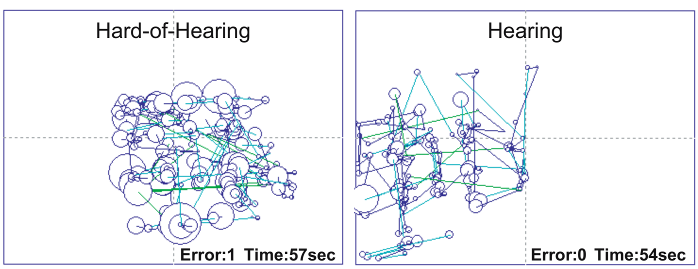


Fig. 1. Differences in scan patterns (adapted from [2])

Number of Errors / Correct Link Selection Times. In order to accomplish the task, the participants had to select a correct link on the top page, but they made a number of wrong selections. Table 1 presents the average number of errors and the average time taken to select the correct link. It is clear that the hard-of-hearing committed errors more often than the hearing did. And the hard-of-hearing took longer than the hearing to select the correct link.

Table 1. Average number of errors and average time to select correct link: hard-of-hearing vs. hearing

	Hard-of-Hearing	Hearing
Average Number of Errors	4.9	2.6
Average Time to Select Correct Link	2 min 42 sec	1 min 27 sec

Processing Times per Link Selection. Figure 2 indicates the time taken to select correct links as a function of the number of link selections. The squares denote the hearing participants. They align on the regression line; the slope corresponds to the time necessary for the hearing to click a link. In contrast, the circles, which denote the hard-of-hearing, do not reveal any correlation. These results imply that the hearing took a certain amount of fixed time before selecting a link, while the hard-of-hearing did not use such a strategic search method when selecting a link.

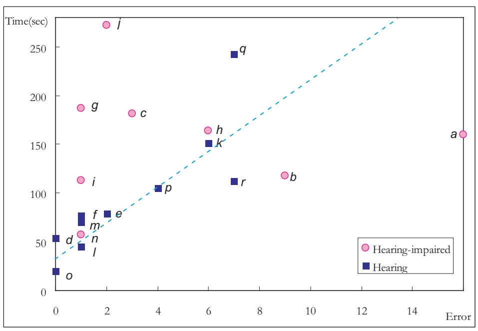


Fig. 2. Time necessary to select a link (adapted from [2])

Nature of Link Selections. We examined the types of the links that the participants selected. The two classes of links are those that contain semantic information and those that contain non-semantic information. The semantic-information links include 1) heading, 2) index, and 3) related (i.e. semantically related to the task goal) links. The non-semantic information links include 1) animation, 2) picture, 3) direct (i.e. cryptic symbol link such as "Z4"), and 4) other clickable objects.

The left pie chart of Figure 3 illustrates the types of links that the hearing participants selected, and the right pie chart depicts those that the hard-of-hearing chose. Overall, the hearing tended to select the semantic-information links more often than the hard-of-hearing did (49% versus 21%). Conversely, the hard-of-hearing selected non-semantic information links more often than the hearing did (62% versus 36%).

Figure 4 combines the two pie charts into a single chart to clarify the differences. The three link types plotted in the upper-right direction correspond

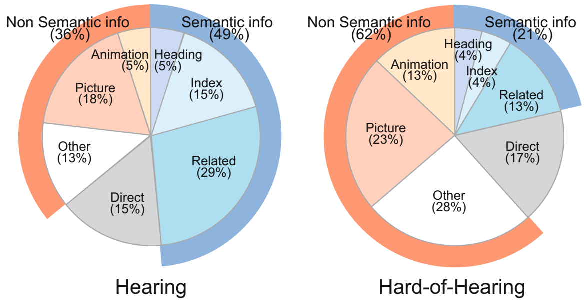


Fig. 3. The types of links that the hearing participants and the hard-of-hearing selected (the pie charts are adapted from [2])

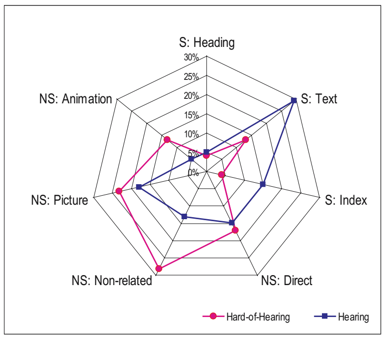


Fig. 4. Differences between the types of links that the hearing and the hard-of-hearing participants selected

to the semantic-information links, and the four link types in the bottom-left direction correspond to the non-semantic information links. The bold line denoting the hearing dominates in the upper-right direction and the thin line denoting the hard-of-hearing dominates in the bottom-left direction. It is clear from the figure that the types of links the participants selected were significantly different in terms of the amount of semantic information contained in the links.

3 Guidelines and Redesign

This section describes the accessibility guidelines that serve as guiding principles for designing accessible Web pages for the hard-of-hearing. It then reports on the redesign of the Web page used for our previous experiment to improve hard-of-hearing usability by following the principles described in the guidelines, and utilizing knowledge about the characteristics of the hard-of-hearing identified in our previous experiments.

3.1 Accessibility Guidelines for Hard-of-Hearing

Designers of Web-based materials should refer to Web design guidelines from W3C's WAI and the US Government Section 508. The guidelines are presented under "Principles," accompanied by an explanation as to who benefits from them.

In W3C's Web Content Accessibility Guidelines [6], Principle 1 suggests that providing alternatives to audio information is the key to Web accessibility for the hard-of-hearing. The US Government Section 508 recommends attaching synchronized captions to audio, video, and multimedia material for hard-of-hearing users. The primary focus of Web materials and computer-based support for the hard-of-hearing is the provision of computer-generated images of sign language and real-time text annotation. Almost all currently used guidelines involve the substitution of audio information as the only aspect of Web-based interaction. These techniques are appropriate for Principle 1. From the viewpoint of Web-material designers, it is technically easy to conform to this principle; we simply need to provide substitutes for auditory information. However, our experiments clearly demonstrated that the hard-of-hearing's style of accessing text information differs from that of hearing persons, and not all hard-of-hearing persons use sign language. Thus, the current use of guidelines for the hard-of-hearing may be seriously limited. W3C's Principle 3, in contrast, indicates that "content and controls must be understandable." This principle is important because Web-based tasks are performed interactively, requiring comprehension of information provided on the computer screen.

3.2 Redesign

We redesigned the page according to Principle 3 by considering interaction characteristics of the hard-of-hearing. More specifically, we conjectured that the design of the experimental Web page was not self-evident from the way the information was organized. Hard-of-hearing participants would have had difficulty capturing hidden semantic structure, partly because their primary language is not written language.

Considering the Web interaction characteristics of the hard-of-hearing, we redesigned the page in two ways: 1) adding vertical lines that should function as visual support enabling the hard-of-hearing to grasp the informational structure easily, and 2) replacing difficult-to-understand labels with comprehensible representations. The left portion of Figure 5 depicts the original design, and the right presents its redesign. We expected that this redesign would improve the site's usability for hard-of-hearing persons. The results of the evaluation confirmed that this redesign was effective.

3.3 Evaluation

Five hard-of-hearing persons participated in the experiment. Their eye movements were recorded with the use of an EMR-HM8 of NAC, Inc. The task images were projected onto a flat screen 150cm in front of the subject. The projection window was 90cm wide by 75cm high, with a viewing angle of 33 degrees horizontal by 27.5 degrees vertical. The data sampling rate was 60 Hz. We recorded the participants' link selections and eye movements, and analyzed the data from the top page. The right portion of Figure 6 depicts a typical eye movement pattern of the hard-of-hearing participants for the redesigned page. Table 2 indicates the improvement of performance in terms of average number of errors and average time taken to select the correct link.

The improvement was significant. The average number of errors decreased from 4.9 for the original page to 0.8 for the redesigned page, and the average time taken to select the correct link decreased to 42 seconds, which was 2 minutes faster than the time required for the original page. The hard-of-hearing's

Fig. 5. Original design and redesign

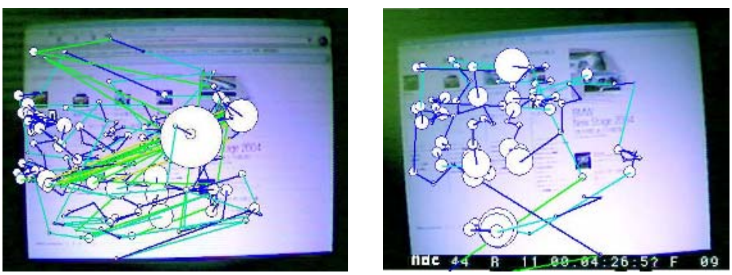


Fig. 6. Scan path of one hard-of-hearing participant (left, original design; right, after redesign)

Table 2. Average number of errors and average time to select correct link of hard-of-hearing: original page vs. redesigned page

	Original Page	Redesigned Page
Average Number of Errors	4.9	0.8
Average Time to Select Correct Link	2 min 42 sec	0 min 42sec

performance on the redesigned page was comparable to or even better than that of the hearing (right column of Table 1). These results imply that, with the redesigned page, the hard-of-hearing participants appropriately captured the informational organization of the page and as a result were able to perform the task efficiently. It was assumed that with the original page the hearing captured the hidden informational structure and therefore performed better than the hard-of-hearing, who had difficulty capturing it. The redesign appeared to eliminate the hard-of-hearing's difficulty in processing ambiguously designed informational structure and to orient the cognitive resources in the right direction.

4 Conclusion and Future Plans

An important lesson is that what is obvious for the Web-literate is not necessarily obvious for the hard-of-hearing. Hidden semantic structures caused by fancy design ideas were not easy for the hard-of-hearing to capture, resulting in a serious usability problem. Their eye movements told us clearly where the source of confusion was and suggested the effective design change. A small design consideration resulted in a large improvement of the Web site's usability.

As stated above, the W3C's Web accessibility guidelines say only *what to do*, but not *how to do* it. This study exemplifies how accessibility can be accomplished. We believe that the accumulation of techniques is important to the achievement of accessibility for the hard-of-hearing. We plan to continue using this approach with various Web sites.

References

1. M. Namatame and M. Kitajima: "Improving web usability for the hard-of-hearing", Proceedings of Eye Tracking Research & Applications Symposium 2006, p. 39 (2006).
2. M. Namatame and M. Kitajima: "Differences in web-interaction styles of hard-of-hearing and hearing persons", Proceedings of the HCI International 2005 (2005).
3. M. Namatame, M. Kitajima, T. Nishioka and F. Fukamauchi: "A preparatory study for designing web-based educational materials for the hearing-impaired", Proceedings of the 9th International Conference on Computers Helping People with Special Needs (ICCHP2004), pp. 1144–1151 (2004).
4. R. Aouf and S. Hansen: "Integration of signage information into the web environment", Proceedings of the 9th International Conference on Computers Helping People with Special Needs (ICCHP2004), pp. 1071–1078 (2004).
5. "`http://elokuvantaju.uiah.fi/2001/english/english.jsp`" (2001).
6. "Web content accessibility guidelines, 2.0 w3c working draft 23 november 2005, `http://www.w3.org/tr/2005/wd-wcag20-20051123/`" (2005).

Accessible Websites for People with Dementia: A Preliminary Investigation into Information Architecture

Nada Savitch and Panayiotis Zaphiris

Centre for HCI Design, City University, London
hf308@soi.city.ac.uk

Abstract. People with dementia have not traditionally been seen as a user group for website development. This paper describes the first attempts to discover some of navigation design needs when developing an information-based website for people with dementia. A card sorting methodology is described using existing information that is provided for people with dementia about their condition. Some participants with dementia found it difficult to group concepts together. This could have a profound affect on the design of good websites.

1 Introduction

There are an estimated 18 million people worldwide with dementia. Dementia primarily affects older people. The chance of having the condition rises with age to 1 person in 20 over the age of 65, and 1 person in 5 over the age of 80 (ADI 2006). There are however maybe 18,000 people under the age of 65 in the UK with dementia (Alzheimer's Society 2006).

People with dementia have not traditionally been seen as a user group for website development. However, due to the availability of drug treatments that slow the progression of the disease and patterns of increasing early diagnosis the numbers of people in the early stages of dementia are increasing. Involvement of people with dementia in voluntary organisations such as the UK Alzheimer's Society is also increasing (Litherland 2004). Websites of organisations for people with dementia now have sections specifically targeted at this user group. However, there is little evidence that the design of these websites takes into account any special needs people with dementia may have when using the web (Savitch and Zaphiris 2005).

Dementia has been defined as a syndrome characterised by the development of multiple cognitive deficits (Cummings and Khachaturian, 1999). Causes of dementia include Alzheimer's disease, vascular dementia and dementia with Lewy bodies (Alzheimer's Society, 2006). The cognitive domain which is impaired first and foremost in Alzheimer's disease is memory (Kertesz and Mohs 1999). Topographical disorientation, ie difficulty in orienting to, navigating through and feeling familiar with one's surrondings is also seen in Alzheimer's disease (Pai and Jacobs 2004). Most people with Alzheimer's disease show language changes with naming disturbances being especially prominent (Henderson 1996). In tests where participants were asked to generate examples of words in a particular category, people with mild and moderate

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 144 – 151, 2006.

Alzheimer's disease produced many fewer exemplars than normal elders. People with mild dementia may remember a given concept and explain its meaning, whereas people with in the moderate stages could not (Bayles 2003). People with Alzheimer's disease may also retain knowledge of low-dominant attributes, ie properties not central to the word's meaning (Nebes and Halligan 1996).

These problems with language are likely to affect the way people with dementia use websites where navigation is heavily based on language and the need to associate the content of a page with a word or phrase used as a menu item or button label. This paper describes the first attempts to discover some of the navigation design needs when developing an information-based website for people with dementia.

2 Involving People with Dementia in Research

Until recently, assumptions were often made that people with dementia do not have views about their condition and are unable to express their own history (Allen et al 2003). However, there is growing recognition that people in the early stages of dementia are able to provide accurate and valid reports of the experience of services provided for them, such as community care (Bamford and Bruce 2000).

When working with people with dementia, appropriate research methodologies are vital. In both the areas of dementia care and dementia research a more person-centred approach has been adopted (Kitwood 1997). A collaborative style is appropriate when carrying out research with this population. In such resarch the person with dementia, carers and researchers explore difficulties and options for their resolution together (Blackman et al 2003). In the HCI community, it is recognised that new techniques might be needed when working with older or disabled people (Newell and Gregor 2004). Older people in general encompass an incredibly diverse group of users. For this reason, Gregor et al (2002) have put forward the concept of user-sensitive inclusive design rather than user-centred.

2.1 People with Dementia and Computers

Some people with dementia, however, are actively using computers and the internet (DASNI 2005; Alzheimer's Forum, 2005). Alzheimer's Forum is a website for people with dementia that is run by a small group of people with dementia at the West Kent branch of the Alzheimer's Society. There are also a number of websites offering information about the condition and ways of coping. However, although the content of these websites is written for people with dementia there is no evidence that they have been designed specifically for this user group (Savitch and Zaphiris 2005).

The aim of the follow-up study, reported in this paper, is to investigate how people with dementia view the information that is available to them in terms of any mental models they may have of how this information should be arranged and organised.

2.2 Information Architecture and Website Design

Information architecture has been defined as the process of creating an underlying organization system for information (Kuniavsky 2003). In website design, it especially refers to how the different pages of the site relate to one another, and involves such issues as content analysis and planning, organisation of the pages, providing cues to help users to orient themselves, labelling and navigation design (Brinck et al 2002). Navigation is a major part of the user experience on a website and careful attention needs to be paid to designing appropriate and usable navigation that allows users to get to where they want to go (Lazar, 2003).

Kuniavsky (2003) has identified the need to establish who the audience is, how they think, what words they use and whether the existing information architecture makes sense to them. Most websites are designed in some sort of top down hierarchy. Usually there is a top page where the user enters the site, middlemen page for topical areas or different user audiences and content pages (Lazar, 2003). Appropriate terminology is important as words can be ambiguous and easily misunderstood and their comprehension by the website user group is critical (Kuniavsky 2003).

The user's mental model of a particular website – ie how people understand the topic, the picture they have developed of how the information is organised, and the names and relationship of the terms they use (Kuniavsky 2003) is investigated in this study using a card sorting technique.

3 Methods

Ten people with dementia were asked to take part in the study. They were all in contact with the UK Alzheimer's Society for support and/or advice and recommended by staff working at the Society. The participants consisted of eight men and two women. Participants were not expected to have any prior knowledge of computers or websites. For comparison the study was also undertaken with eight people who are employed by the Society in various roles to give out information about dementia. This group consisted of six women and two men. This group was chosen because they are instrumental in providing information for people with dementia. Assumptions are often made that this group will have insight into how people with dementia think.

The topics used in this study were taken from existing printed publications produced by the UK Alzheimer's Society. The information provided and the language used has been extensively researched by the Alzheimer's Society 'Living with dementia' programme to produce a booklet entitled 'I'm told I have dementia' and a series of information sheets for people with dementia. This information has been used to produce a section of the UK Alzheimer's Society's website (Table 1). The website section was developed by the Society's website manager using a 'bottom-up' approach (Brinck et al 2002) using the text and titles of the sheets and booklet chapters. It has been established that the existing navigation is not particularly easy for people with dementia to use (Savitch and Zaphiris 2005).

3.1 Card Sorting

The card sorting technique is a useful approach to understanding what natural categories people have for the domain, and is appropriate when the designer is not a domain expert and needs the insight of users (Brinck et al 2002). The card sorting technique was chosen for this study because it is a simple technique that is not threatening to the participant. Card sorting has successfully been use to investigate web health information architecture for older users (Kurniawan and Zaphiris 2003).

The titles of the sheets/booklet chapters were cut out and pasted onto cards. Participants were asked to 'sort' the cards into groups in any way that made sense to them. It was made clear to all participants that there was no right or wrong answer, and that the groups could be made up of any number of cards and that not all the cards need to be grouped. Once the cards were sorted, participants were asked to provide a suitable overall title for each group. Participants were encouraged to talk through their decisions throughout. The results were then analysed using 'eyeballing' and cluster analysis techniques (Martin 1999).

Table 1. Existing website design. An additional card - 'Having a brain scan' was also used.

Website heading	Card topic
After the diagnosis	Living with dementia
Types of dementia	What is dementia? What is Alzheimer's disease? What is vascular dementia? What is dementia with Lewy bodies? What is fronto-temporal dementia?
Treatments	Aricept – a drug for Alzheimer's disease Reminyl – a drug for Alzheimer's disease Exelon – a drug for Alzheimer's disease Ebixa – a drug for Alzheimer's disease
Where can I get support?	Where can I get support? How can the GP help? How can social services help?
Planning for the future	Sorting out your money Working Will I still be able to drive? Writing a will
Ideas for looking after yourself	Everyday tips Making everyday life easier Ideas for keeping healthy
Family and friends	Telling other people Will my family get Alzheimer's disease?

4 Results

Three of the ten participants with dementia did not group the cards at all. They were happy to talk about the individual topics raised, but did not find any associations between the topics. All three participants could read the topics and clearly understood the meaning of the word or phrases. They appeared however to sometimes be confused especially if the phrase on the card appear in the form of a question. This would often prompt the participant to answer the question or state that they knew the answer.

The results of the card sort for the remaining seven participants with dementia and the eight information workers were analysed with the IBM's EZSort software. The results can be seen in Figure 1. The aaverage algorithm was used. This gives a balance between two algorithms - single linkage algorithm, which emphasises more on similarities and the complete linkage algorithm, which emphasises more on differences.

The analysis of the card sorts shows that all participants in the study have grouped the drugs for Alzheimer's disease together and the various causes of dementia together. From Figure 1, it is clear that there is more agreement in associations within the group of information workers than within the group of people with dementia. The results from the people with dementia are less uniform than those of the information workers.

4.1 Areas of Most Difference Between the Sets of Participants

The topics that showed the most difference between the two sets of participants were: 'Telling other people', 'How can my GP help?' and 'Living with dementia'. The participants with dementia generally kept 'Telling other people' as a topic on its own. Many of the participants highlighted this subject as being very important to them. The information workers were more likely to group this card with others. The topic 'How can my GP help?' was treated differently by the two sets of participants. The information workers tended to see this topic as completely separate from the 'Seeing your doctor' topic and tended to associate it with the general 'How do I get support' topic and the 'How can social services help?' topic. However, the participants with dementia generally kept the two topics 'Seeing your doctor' and 'How can my GP help?' together. Presumably because they did not differentiate between the two statements in the same way as the information workers did. The 'Living with dementia' topic was also viewed differently between the two groups. The information workers tended to see this as a general topic covering everyday life for someone with dementia. Whereas the people with dementia themselves seem to associate it more with the more practical and pressing issues such as driving and working.

4.2 Areas of Most Agreement Between the Sets of Participants

The two areas where there was most agreement between the two sets of participants were the groups of topics associated with drug treatments and the group of topics associated with the different causes of dementia. Both sets of participants also tended to keep the 'Will my family get Alzheimer's' topic on its own. The information workers were more likely to name the drugs group of cards 'Treatments' (4/8 participants), whereas the people with dementia were more likely to call the group 'Drugs' or 'Medicines' (4/7 participants). The existing website uses the word 'Treatments'. The participants with dementia each called the causes of dementia group a different title, eg what is dementia, types of dementia, forms of dementia, kinds of dementia, whereas four of the eight information workers chose the title 'Types of dementia'. Again, 'Types of dementia' is the phrase used on the existing website. The 'Will my family get Alzheimer's' card was classed as a topic on its own by three of the seven people with dementia, and two of the eight information workers. One person with dementia and two out of eight information workers grouped it with other 'medical' cards. Two of the information workers grouped the card with the 'Telling other people' card. Again, this might reflect their previous knowledge of the existing website.

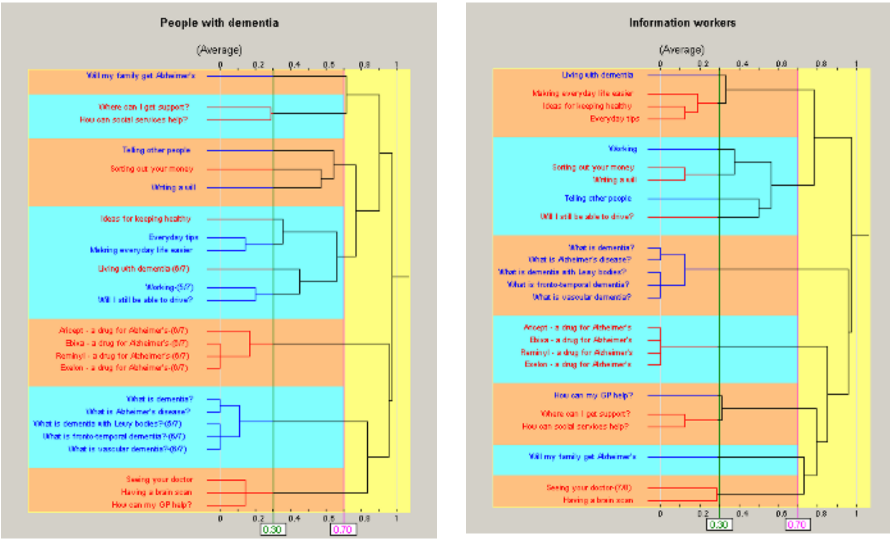


Fig. 1. Cluster analysis of card sorting data

4.3 Number, Size and Names of Groups

The information workers seem to have sorted the cards into fewer groups – 4-11 groups whereas the people with dementia had 5-9. People with dementia tended to have more groups consisting of only one card (14) than the information workers (12). Only three out of the eight information workers had any groups with only one card, whereas six out of seven people with dementia chose to have at least one group consisting of a single card. Although both the controls and the group of people with dementia found it difficult to give names to the groups, the people with dementia found it more difficult. Often the names of the groups given by people with dementia were longer and more complicated that those given by the control group. The group names given by people with dementia had a mean of 3.8 words per title compares with 3.0 words for the information workers.

4.4 Carrying Out the Task

Way in which the task was carried out was very similar. All participants tended to pick out the cards about drug treatments and the types of dementia first. Some participants with dementia were put off by terms that they didn't understand, eg some seemed confused about drugs names or the causes of dementia. Most of the participants with dementia were keen to give their opinions about the topics but were easily distracted – especially in the topic was given in the form of a question.

5 Conclusions and Discussion

The finding that three of the ten participants with dementia did not group the cards at all is interesting. Although it could be that the participants did not understand the task that was being asked of them. The fact that people with dementia found it difficult to group concepts together could have a profound affect on the design of good websites. It could be argued that people in a stage of dementia that hinders them from thinking hierarchically would not be able to use the world wide web. However, the authors believe that future generations of people with dementia will turn to the web for information and support. It is the role of web designers to make the website as easy to use as possible. It might be that traditional website design based on menus hierarchies is not suitable for people with dementia. However, it is too early in the research to make generalizations about website design.

The fact that three of the participants had problems understanding the task raises the issue that the card sorting technique used may not be suitable for use with people with dementia. It might be tempting to dismiss the views of the three participants who did not group the cards because there might be an assumption that these people would not be able to use websites. However, there is no evidence that the participants could not use a computer or that they would not be interested in the information provided by the Alzheimer's Society website. More work is needed to establish how people with dementia do find information they are looking for.

The interviews with people with dementia were generally longer than those with the information workers. All the people with dementia were keen to discuss the topics on the cards and their experiences. This was especially true when the topic was written as a question.

The fact that all participants grouped the cards relating to drug treatments and the causes of dementia may be due to the topics but may also be because the terminology used is very similar. For example, all the cards about drug treatments are written 'Drug name – a drug for Alzheimer's' and all the cards on the topic of causes of dementia are written in the style 'What is xxx?'. Terminology appears to be important – the information workers clearly see a difference between the phrase 'seeing your doctor' and the phrase 'how can my GP help?', whereas the group of people with dementia linked these two cards together more. The issue of the use of questions as headings was also raised. One of the information workers was initially drawn to grouping the cards as 'questions' and others. One of the participants with dementia specifically highlighted the issue of whether the 'reader' was 'you' or 'me'.

Further research is need. It is proposed that a website be devised based on the grouping identified by people with dementia and tested by people with dementia. The terminology used in website design also needs to be investigated further. Navigation systems are clearly important for people with dementia. The idea of a flat structure needs to be investigated further and compared with other navigation systems. It is clear that traditional HCI methodologies need to be adapted when designing for people with dementia.

References

1. Allen C, Newby G and Kennally M (2003). Forthcoming research in the Oxford region: innnovations in dementia care – the benefits of multi-media profiling for older people with dementia, PSIGE Newsletter, 83 May 2003, 13-14.
2. ADI - Alzheimer's Disease International (2006). `http://www.alz.co.uk`
3. Alzheimer's Forum (2006). `http://www.alzheimersforum.org`

4. Alzheimer's Society (2006). http://www.alzheimers.org.uk
5. Bamford C and Bruce E (2000). Defining the outcomes of community care: the perspectives of older people with dementia and their carers. Ageing and Society, 20, 543-570.
6. Bayles, K A (2003). Effects of working memory deficits on the communicative functioning of Alzheimer's dementia patients. Journal of Communication disorders, 2003, 36, 209-219.
7. Blackman T, Mitchell L, Burton E, Jenks M, Parsons M, Raman S and Williams K (2003). The accessibility of public spaces for people with dementia: a new priority for the 'open city', Disability and Society, 18 (3), 357-371.
8. Brinck T, Gergel D, Wood S C (2002). Designing websites what work: usability for the web. Morgan Kaufmann, San Francisco.
9. Cummings J L and Khachaturian Z S (1999). Definitions and diagnostic criteria. Chapter 1 in Clinical diagnosis and management of Alzheimer's disease, 2nd ed. Ed Serge Gauthier. 1999. Martin Dunitz, London.
10. DASNI (2005) http://www.dasni.org
11. Gregor P, Newell A F, Zajicek M (2002). Designing for dynamic diversity – interfaces for older people, ASSETS 2002. The 5th International ACM conference on Assistive Technologies, Edinburgh, Scotland, 8-10 July 2002, pp151-156.
12. Henderson V W (1996). The investigation of lexical semantic representation in Alzheimer's disease, Brain and Language, 54, 179-183.
13. Kertesz A and Mohs R C (1999). Cognition. Chapter 12 in Clinical diagnosis and management of Alzheimer's disease, 2nd ed. Ed Serge Gauthier. 1999. Martin Dunitz, London.
14. Kitwood T (1997). The experience of dementia. Aging and Mental Health, 1 (1), 13-22.
15. Kuniavsky M (2003). The user experience. Chapter 4 in Observing the user experience: a practitioner's guide to user research. Morgan Kaufmann Publishers, San Francisco, pp43-53.
16. Kurniawan S H and Zaphiris P (2003). Web health information architecture for older users, IT & Society, 1 (3), 42-63.
17. Lazar J (2003). The World Wide Web. Chapter 36 In: The Human-Computer Interaction Handbook: fundamentals, evolving technologies and emerging applications (ed Julie Jacko and Andrew Sears). Lawerence Erblaum Associates, New Jersey, pp714-730.
18. Litherland R (2004). 'Listen to us', Working with Older People, 7 (4), 17-20.
19. Martin S (1999). Cluster analysis for website organization, Internetworking, Dec 1999.
20. Nebes R D and Halligan E M (1996). Sentence context influences the interpretation of word meaning by Alzheimer's patients. Brain and Language, 54, 233-245.
21. Newell AF and Gregor P (2004). CHI and older and disabled people, Proceedings of HCI 2004: Design for life, Leeds, Volume 2, pp205-206.
22. Pai Ming-Chey and Jacobs W J (2004) Topographical disorientation in community-residing patients with Alzheimer's disease, International Journal of Geriatric Psychiatry, 19: 250-255.
23. Savitch N and Zaphiris P (2005) An investigation into the accessibility of web-based information for people with dementia, 11th International Conference on Human-Computer Interaction, Las Vegas, July 2005.

People with Disabilities: Automatic and Manual Evaluation of Web Sites

Introduction to the Special Thematic Session

Helen Petrie and Gerhard Weber

University of York, Department of Computer Science
Heslington, York, YO10 5DD, UK
Helen.Petrie@cs.york.ac.uk
Universität Kiel, Multimedia Campus, Westring 431-451, 24118 Kiel
g.weber@mmc-kiel.com

Abstract. Quality of websites is a key factor in addressing all users. Besides functional issues is the design of each web page affecting the ability to navigate and interact with web applications. Web designers are only slowly becoming aware of accessibility issues, more tools, better processes in creating high quality websites and a better understanding of guidelines is needed.

1 Introduction

Surveys have repeatedly shown that the accessibility of websites for people with disabilities is disappointingly low. In 2004 the Disability Rights Commission [1] in Great Britain found that only 19% of a very large and diverse sample of websites passed even the most basic accessibility criterion, automated testing for Priority 1 level checkpoints of the Web Content Accessibility Guidelines (version 1, WCAG1)[1]. In 2005, a survey of public administration websites across Europe[2] found that only 10% passed the automated Priority 1 tests, and only 3% passed the full set of Priority 1 tests (achieved with both automated and manual testing). Smaller surveys and surveys of more specific web domains, both in Europe and other parts of the world, show similarly low levels of conformance to WCAG [2 – 7]. Two recent historical studies also paint a similar picture: Hackett, Parmanto & Zeng [8] found that a random selection of websites (accessed using the Internet Archive's Wayback Machine) significantly decreased in accessibility during the period 1997 to 2002; Petrie [9] in a meta-analysis of 11 studies conducted between 1999 and 2005 found no significant change in the overall level of accessibility of websites.

This situation is extremely disappointing, given legislation in many countries relevant to web accessibility and the many initiatives to promote web accessibility. It is also surprising, given the range of tools available to web authors to assist them when developing accessible content for the web. Such tools, all of which are useful to web authors, include:

1 http://www.w3.org/TR/WCAG10/
2 www.cabinetoffice.gov.uk/e-government/docs/eu_accessibility/pdf/eaccessibility(eu)_report.pdf

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 152 – 155, 2006.

- authoring tools that provide guidance on accessibility (e.g. Dreamweaver[3] provides accessibility dialogue boxes[4] that prompt the author to enter accessibility tags and attributes),
- tools that can be used to check for specific accessibility issues although they were not designed for this purpose (e.g. the Lynx browser[5] allows an author to view the linearized version of a web resource, thus gaining an idea of how it will be read to blind web users via a screenreader[6]),
- tools that are developed to visualize specific accessibility issues (e.g. there are a number of tools that attempt to show how a web resource will appear to users with different forms of colour vision deficiency),
- tools that provide easy access to a range of specific checking capabilities, such as those discussed above, e.g. the Web Accessibility Toolbar[7],
- automated evaluation and evaluation and repair (E&R) tools, that assess the conformance to some of the WCAG checkpoints, and may assist the author in repairing violations of the checkpoints[8],
- testing web resources with assistive technologies, such as screenreaders for blind users and software for dyslexic users, to check how they are rendered in these technologies, and
- testing web resources with disabled web users to ensure that these groups can easily use the resources.

This Special Thematic Session (STS) addresses some of the important issues currently challenging web developers and website owners in relation to the evaluation of websites for accessibility, using both manual and automatic testing. Authors of the papers address current problems around the processes associated with both small scale and large scale accessibility testing and refer increasingly to the multiplicity of definitions of accessibility. For example, there are differences in the level of abstraction of criteria in different definitions (which implies also the applicability of automatic testing tools), there are large variations among definitions with respect to their ability to compare different levels of accessibility and there are differences in the legal definitions in different countries. In our STS, Brajnik [10] particularly provides an interesting discussion of the problems of defining web accessibility.

Since the development of WCAG 1.0, it is clear that the criteria for web accessibility will need to be periodically revised in order to keep them up-to-date and improve them where possible. Updates and improvements can arise from better assistive devices, better implementations of browsers, new technologies supporting browsers such as mobile phones, or better understanding of the evaluation of some

---

[3] Examples of commercial products are provided only as illustrative of particular points, and are not endorsements of these products.

[4] http : //livedocs.macromedia.com/dreamweaver/8/using/wwhelp/wwhimpl/common/html/wwhelp.htm?context=LiveDocs_Parts&file=15_imag4.htm#71908

[5] http://lynx.browser.org/

[6] For example, Vischeck (http://www.vischeck.com/)

[7] http://www.visionaustralia.org.au/ais/toolbar/

[8] For a comprehensive list, see http://www.w3.org/WAI/ER/existingtools.html

criteria. An approach to define the criteria more precisely and also make them suitable for use by automatic testing tools is discussed by Strobbe et al [11], who propose a Test Case Description Language (TCDL). The authors demonstrate a set of test materials and a test procedure in order to validate the accessibility criteria and also to compare automatic tools using the criteria. Moreover, they expect their TCDL can also be used for novel technologies such as XForms and Scalable Vector Graphics (SVG) as the appropriate browsers and criteria for accessibility of these technologies become available. Leporini et al [12] also address the need to handle multiple guidelines and propose an executable description of guidelines. This paper demonstrates their language for some selected examples and reports on a run-time environment for executing guideline descriptions. Abou-Zahra proposes in another paper [13] the use of the Evaluation and Report Language (EARL) based on semantic web technologies to allow a comparison among such test results.

The validation of guidelines can increase our understanding of limitations of guidelines. Herramhof et al [14] report about the development of software tools to support the authors of evaluation criteria and a tool for conducting remote evaluations with heterogeneous user groups or automatic testing tools in parallel. They expect current attempts to validate guidelines may to lead to more confidence in the quality of automatic testing results and help to reduce the need for expert and user testing.

Some guidelines still lack appropriate automatic tools to evaluate them. These include the criteria of "making the text content readable" and "mechanism to find context [… and navigate through it" [15]. Two papers in the STS address each of these guidelines separately and report on ongoing work. Findings from computer linguistics can be applied in order to help to measure the simplicity of language on the web. Monitoring the actual navigation by end users within and between web pages is a step towards better analysis of the mechanisms offered for navigation. Gonzalez et al [16] propose implementing agents to better monitor user navigation when conducting remote user testing. This will allow the differences between different assistive technologies to be taken more automatically into account.

Two issues of evaluating accessibility still require either automatic, expert or end user testing: periodic assessment of very large number of web sites for benchmarking purposes and task-driven expert or usability testing. Several papers in the STS [17, 18] report on ongoing work towards large scale testing and have to tackle the temporal nature of dynamic web sites. Only by regularly testing of large numbers of websites can an analysis of the trends and problems of web accessibility be undertaken.

Tools for the integration of automatic testing with user and manual testing are still in their initial states of development. Rangin [19] proposes to ensure use of best practices and apply multiple and different tools. Benavídez et al [20] discuss their approach to semi-automated testing and address the requirements for next generation of testing tools: The heterogeneity of users with different requirements is not yet supported by either automatic tool or tools for manual testing. In addition, in order to make testing both more effective and efficient (and hence economic and attractive to web developers and owners), the workflow associated with testing by automated tools, experts and end users needs to be supported better by all existing tools.

References

1. Disability Rights Commission (2004). *The web: Access and inclusion for disabled people*. London: The Stationery Office.
2. Flowers, C.F., Bray, M. and Algonzzine, R. (2000). Accessibility of Schools and Colleges of Education Home Pages for Students with Disabilities. *College Student Journal*. 34(4), 550-556.
3. Kelly, B. (2002). WebWatch: An accessibility analysis of UK university entry points. *Ariadne*, 33. Available at: http://www.ariadne.ac.uk/issue33/web-watch
4. Kelly, B. (2004). Accessibility Survey of UK University Entry Points. *Institutional Web Management Workshop 2004*, University of Birmingham. Available at: http://www.ukoln.ac.uk/web-focus/events/workshops/webmaster-2004/talks/phipps-kelly/survey/.
5. Loiacono, E.T. (2004). Cyberaccess: Web accessibility and corporate America. Communications of the ACM, 47(12), 83 – 87.
6. Schmetzke, A. (2001). Web accessibility at university libraries and library schools. *Library Hi Tech,* 19(1), 35 – 49.
7. Sullivan, T. and Matson, R. (2000). Barriers to Use: Usability and Content Accessibility on the Web's Most Popular Sites. *Proceedings of the Conference of Universal Usability*. Arlington, VA: ACM.
8. Hackett, S., Parmanto, B. and Zeng, X. (2004). Accessibility of internet websites through time. Proc. ASSETS 2004. New York: ACM Press.
9. Petrie, H. (2005). Recent developments in website accessibility. Keynote paper at Conference on Accessible Design in the Digital World. Dundee, Scotland.
10. Brajnik, G (2006). Web accessibility testing: when the method is the culprit. This volume.
11. Strobbe, C., Herramhof, S., Vlachogiannis, E. and Velasco, C.A. (2006). Test case description language (TCDL): Test case metadata for conformance evaluation. This volume.
12. Leporini, B., Paterno, F. and Scorcia, A. (2006). An environment for defining and handling guidelines for the web. This volume.
13. Abou-Zahra, S. (2006). Managing and monitoring web site accessibility. This volume.
14. Herramhof, S., Petrie, H., Strobbe, C., Vlachogiannis, E., Weimann, K., Weber, G. and Velasco, C. (2006). Test case management tools for accessibility testing. This volume.
15. Caldwell, B., Chisholm, W., Slatin, J., Vanderheiden, G., and White, J. (2005). Web Content Accessibility Guidelines 2.0, W3C Working Draft 30 June 2005. World Wide Web Consortium. Available at: http://www.w3.org/TR/WCAG20/
16. Gonzalez, M., Gonzalez, M., Rivera, C., Pintado, I. and Vidau, A. (2006). Testing web navigation for all: an agent-based approach. This volume.
17. Snaprud, M.H., Ultveit-Moe, N., Pillai, A.B. and Olsen, M.G. (2006). A proposed architecture for large scale web accessibility assessment. This volume.
18. Petrie, H. …. (2006). …. This volume.
19. Rangin, H. B. (2006). University of Illinois tools and techniques for functional web accessibility. This volume.
20. Benavídez, C., Fuertes, J. L., Gutiérrez, E., and Martínez, L. (2006). Semi-Automatic Evaluation of Web Accessibility with HERA 2.0. This volume.

Web Accessibility Testing: When the Method Is the Culprit

Giorgio Brajnik

Dip. di Matematica e Informatica
Università di Udine
Italy
www.dimi.uniud.it/giorgio

Abstract. Testing accessibility of a web site is still an art. Lack of appropriate definitions of accessibility and of standard testing methods are some of the reasons why Web accessibility is so difficult to achieve.

The paper describes a heuristic walkthrough method based on barriers; it then discusses how methods like this can be evaluated, and it shows experimental data about validity and usefulness of the method when compared to *standards review.*

1 Introduction

Unless one has a clear notion of the property that has to be tested, the testing process is rarely successful. And when the property depends on human cognitive processes, then also the testing method may reduce the effectiveness of testing. This is indeed the case for web accessibility. On the one hand, there are several (more or less practical) definitions of accessibility. Sometimes accessibility is defined in terms of effectiveness; now and then it is defined in terms of usability; but unfortunately there are too often claims that a web site is accessible simply because an automatic testing tool yielded no error.

While accessibility can be tested based on guidelines (like WCAG 1.0, or Section 508) through a *standards review* method, other methods can be employed, like user testing [4] or usability inspection methods [14,7,12] or those suggested in [8]. Knowledge about the benefits and shortcomings of these methods applied to accessibility is still missing. If it were available, then methods appropriate to the specific situation at hand could be used, and results could be easily compared (over time for the same web site and type of users, or between different web sites or populations with differing characteristics).

To be really useful, evaluation methods should constrain the way in which the evaluator identifies problems and how they are graded in terms of importance. Only when these two kinds of decisions can be standardized, then the results produced can be used to rank web sites and to prioritize their bugs. Prioritization of defects is of paramount importance as any web developer required to fix them works always in a *scarce resource* mode.

The purpose of this paper is to propose and evaluate a heuristic walkthrough method to fill what I consider to be a gap in evaluation methods for accessibility

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 156–163, 2006.

assessments. Heuristics are interpretations and extensions of well known accessibility principles; within this method they are linked to user characteristics, user activities, and situation patterns so that appropriate conclusions about user effectiveness, productivity, satisfaction and safety can be drawn, and appropriate severity scores can be consequently derived.

2 Accessibility Testing

2.1 Problems in Accessibility Testing

As discussed in [1] at least three definitions of accessibility exist: some refer to usability, some to effectiveness[1] and some to other principles like perceivability, understandability and operability. The problem is that depending on the definition different methods have to be used to investigate. Running a user test doesn't make sense if we want to determine conformance to guidelines, and conversely conformance testing cannot be used to determine usability of the web site with respect to disabled users.

More confusion exists regarding the methods to use. For example, the current Italian regulation for web accessibility [13,6] specifies a number of technical requirements similar to WCAG 1.0 and Section 508 points. However, in order to achieve a certification mark evaluators have to perform a *cognitive walkthrough* (an analytical method for early–on usability investigations, that is yet unproven as a method for accessibility testing). In addition, the regulation specifies 12 general usability principles that are generally employed with a different method, namely *heuristic evaluation*. It also requires that evaluators classify identified problems into 5 severity levels, without specifying how severity should be determined. It then suggests empirical methods that have no proved effectiveness (*e.g. subjective assessments*) and finally it requires that evaluators compute a final score for the site on the basis of mean averages of severity levels associated to problems (an ineffective aggregation technique). Although such a regulation sets a certification framework for web accessibility, in my view it is unlikely to succeed because of extreme subjectivity and variability, poor practicality and measure–theoretical shortcomings.

Recently, other initiatives towards accessibility certifications took place [5]: but unless adoption of proper methods is prescribed, the desired connection between quality marks and actual accessibility levels will be missing, consistently with those that argue against any accessibility certification program [10].

Even simpler methods like *conformance testing* (called also *standards review*) are not problem–free. As discussed in [11], WCAG 1.0 guidelines suffer from their theoretical nature, dependency on other guidelines, ambiguity, complexity, their closed nature and by some logical flaws. Most of these comments apply as well to the current WCAG 2.0 draft, to Section 508 and to the Italian official

[1] In the following I assume that accessibility is defined as: "*... web sites are accessible when individuals with disabilities can access and use them as effectively as people who don't have disabilities*" [15, p. 3].

technical requirements. Guidelines are intentionally defined in terms that are independent from the technology used in implementing and in visiting web sites. As a consequence, guidelines are often too abstract to be directly applicable to a web site, creating a gap that has to be filled by the evaluator. In addition, they don't help the evaluator in distinguishing important problems from trivial ones. For example, few of the images in a web site that lack an appropriate alternative text are a true barrier: most of the images are used for emotional purposes, which in textual alternatives would almost always be lost anyway. But an important function asked to an evaluator is to tell what the consequences of such defects are: this, however, can be done only if appropriate use scenarios are considered. The *standards review* method does not tell how to choose scenarios and how to rate the defect, except for static priorities that cannot reflect these scenarios.

2.2 Requirements on Methods

Evaluation methods should be valid, reliable, useful and efficient. *Validity* is "the extent to which the problems detected during an evaluation are also those that show up during real–world use of the system" whereas *reliability* is "the extent to which independent evaluations produce the same result" [7,9]. *Usefulness* is the effectiveness and usability of the results produced (with respect to users that have to assess, or to fix, or otherwise to manage accessibility of a web site). Finally, *efficiency* is given by the resources being utilized during an evaluation (in terms of time, persons, skill level, facilities, money) related to some level of validity, reliability and/or usefulness.

Notice that especially validity and reliability are negatively affected by a number of factors: different evaluators may have different levels of knowledge in web accessibility, in assistive technology, in the relevant standards, in the behavior of browsers. Hence it is important that the method constrains as much as possible the subjectivity of the evaluators in identifying failure modes[2], in diagnosing them and in judging them.

3 Barriers Walkthrough

An *accessibility barrier* is any condition that makes it difficult for people to achieve a goal when using the web site through specified assistive technology (see figure 1 for an example). A barrier is a failure mode of the web site, described in terms of (i) the user category involved, (ii) the type of assistive technology being used, (iii) the goal that is being hindered, (iv) the features of the pages that raise the barrier, and (v) further effects of the barrier.

Barriers to be considered are derived by interpretation of relevant guidelines and principles [4,16]. A complete list can be found in [2].

[2] I use the term *failure mode* to mean the way in which the interaction fails; the *defect* is the reason for the failure, its cause; the *effects* are consequences of the failure mode.

To apply the *barriers walkthrough* method (BW), a number of different scenarios need to be identified. A scenario is defined by user characteristics, settings, goals, and possibly tasks of users belonging to given categories. At least categories involving blind users of screen readers, low–vision users of screen magnifiers, motor–disabled users of a normal keyboard and/or mouse, deaf users, and cognitively disabled users (with reading and learning disabilities and/or attention deficit disorders) should be considered.

barrier	users cannot perceive nor understand the information conveyed by an information rich image (*e.g.* a diagram, a histogram)
defect	an image that does not have accompanying text (as an ALT attribute, content of the OBJECT tag, as running text close to the picture or as a linked separate page)
users affected	blind users of screen readers, users of small devices
consequences	users try to look around for more explanations, they spend substantial time and effort; effectiveness, productivity, satisfaction are severely affected

Fig. 1. Example of barrier

User goals and tasks can be defined only with reference to the site being tested. For a web application, one should consider some of the possible goals and tasks usually documented in *use cases* and cross these goals with user categories to obtain the relevant scenarios. For information web sites, a sample of possible information needs can be considered and crossed with user categories. In this way, most of the times, each user goal/task will be associated to different sets of pages to test, and these will be crossed to user categories.

Evaluators then analyze these pages by investigating the presence of barriers that are relevant to the particular user category involved in the scenario.

Cross–checking a barrier to a set of pages in the context of a scenario enables evaluators to understand the impact that the barrier has with respect to the user goal and how often that barrier shows up when those users try to achieve the goal.

In the end, evaluators produce a list of problems, associating each problem to a barrier showing up in a given scenario, to a severity level, and possibly to performance attributes that are affected (*e.g.* effectiveness, productivity, satisfaction, safety).

4 Evaluating Testing Methods

Evaluations of a testing method should determine its validity, reliability, usefulness and/or efficiency. In this paper I report an experimental evaluation of validity and to some extent of usefulness of the BW method compared to conformance test (CT).

Validity and usefulness are determined by computing certain metrics on reports[3] written by 3^{rd} year university students doing their term projects in a class on web design. A comparative study was set up between reports produced using the BW method and reports of conformance tests (CT) based on WCAG 1.0 AA checkpoints. Nineteen different reports produced by 8 different student teams were analyzed; 8 reports were based on conformance test and 11 on barriers walkthrough. These reports were about 6 different public web sites; each web site was evaluated by at least 2 teams using both methods.

The metrics are based on whether a problem is a true or false one, and whether there are false negatives (*i.e.* problems missed by the report). In order to determine the true problems of a web site, the instructor[4] took all the reports regarding a web site and marked out the issues that are wrongly raised (*e.g.* students not having properly understood a checkpoint or barrier). Repetitions of the same problems were not counted, leading to a list of unique (true or false) problems. For reports produced by the CT method, that don't have the severity score, a new score was added by the instructor: when the issue is mentioned in the executive summary of the report it gets a severity score of 3, otherwise 1. A further severity score was added for each issue, representing the severity of the issue that the instructor deemed appropriate. Such a score is 0 for issues that are not a problem, and 1 to 3 otherwise.

Each issue therefore consists of the tuple (team, site, type of problem, method, team–severity, instructor–severity). Data were then aggregated so that at most a single tuple results for any combination of site, team, method and problem type. Severities were aggregated by computing the maximum (*i.e.* by deriving the highest value for the aggregated issues, which corresponds the worst case).

The set of true problems for a web site is then the union of all the problem types that were labeled as "true" by the instructor.

Comparison metrics include:

- *precision* (P), the percentage of reported problems that are true problems; in the following, P is precision, whereas P_3 is the percentage of reported problems that have an instructor–assigned severity of 3;
- *sensitivity* (S), the percentage of the true problems being reported (notice that P and S are totally independent; for example, $P = 1$ also in the extreme case where only 1 true problem was reported, but the web site contains many other unreported problems); S_3 is the sensitivity with respect to the problems having instructor–assigned severity of 3;
- *fallout* (F), the percentage of false problems that are reported;
- *E–measure* (E): a combination of precision and sensitivity in the range $[0, 1)$: $E = PS/(\alpha P + (1 - \alpha)S)$; I set $\alpha = 0.5$ and then $E = 2PS/(P + S)$; E is a monotonic and symmetric function on both arguments; E_3 is the combination of P_3 and S_3;

[3] Available on–line at `www.dimi.uniud.it/giorgio/dida/psw/galleria/galleria.html`. The experimental data used in this paper is available online at [2]. The BW method description that students followed is [3].

[4] And author of this paper.

- *mean severity* ($\overline{sev}$): the mean of the severity assigned by students to true problems;
- *n. of problems with severity=3*: the number of true problems associated to the highest severity level.

Precision, sensitivity, fallout and e–measure are related to method validity; mean severity, the number of high severity true problems, P_3, S_3, and E_3 are related to the usefulness of the method (*i.e.* its ability to focus the evaluator resources onto most important problems).

4.1 Results

Collectively the reports raised 303 problems (respectively 166 for CT and 137 for BW), with 260 true problems (86%).

Figure 2 shows the validity and usefulness metrics split by web site.

calabria	**BW**	**CT**
P	0.85	0.80
P_3	0.08	0
S	**0.48**	**0.70**
S_3	1	0
F	0.33	0.66
E	**0.61**	**0.74**
E_3	0.14	0
$\overline{sev}$	1.91	1
n sev=3	3	0

campania	**BW**	**CT**
P	1	0.76
P_3	0.24	0.12
S	0.75	0.46
S_3	0.83	0.33
F	0	1
E	0.86	0.58
E_3	0.37	0.17
$\overline{sev}$	1.69	1.62
n sev=3	6	4

fvg	**BW**	**CT**
P	1	0.69
P_3	0.20	0
S	**0.29**	**0.85**
S_3	1	0
F	0	1
E	**0.45**	**0.76**
E_3	0.33	—
$\overline{sev}$	2.2	1
n sev=3	4	0

molise	**BW**	**CT**
P	1	0.73
P_3	0.36	0
S	0.74	0.42
S_3	1	0
F	0	1
E	0.85	0.53
E_3	0.53	0
$\overline{sev}$	2.43	2.25
n sev=3	8	5

puglia	**BW**	**CT**
P	1	0.81
P_3	0.60	0.41
S	**0.57**	**0.63**
S_3	0.63	0.58
F	0	1
E	0.73	0.71
E_3	0.62	0.48
$\overline{sev}$	2.34	1
n sev=3	12	0

toscana	**BW**	**CT**
P	0.95	0.84
P_3	0.38	0.03
S	**0.56**	**0.72**
S_3	0.89	0.11
F	0.11	0.89
E	**0.70**	**0.78**
E_3	0.53	0.05
$\overline{sev}$	2.10	1.17
n sev=3	11	3

Fig. 2. Results for the validity and usefulness metrics split by web site; $\overline{sev}$ is the mean severity. Values highlighted in boldface are the only ones where CT scores better than BW. Notice that for each web site the sum of fallout is always 1 since there are only two methods being considered that can produce false positives.

In most of the cases BW is more valid and useful than CT. In fact:

Precision. In 4 cases out of 6, precision for BW is absolute (100%); it is never less than that of CT, never smaller than 85% and it is between 5 to 45%

higher than that obtained through CT. This means that while CT produces between 31 and 16% of false positives, BW yields 15% at the most.

When restricted to severity 3 problems, precision of BW drops, but it never goes below that of CT.

Sensitivity. In 4 cases BW yields a sensitivity lower (i.e. worse) than CT; the minimum for BW is 29% while for CT it is 42%. This means that BW yields a smaller proportion of true problems, which is explainable because some guidelines do not correspond to any barrier in the scenarios that were considered by students. On the other hand, by reducing the analysis to true problems with severity 3 then in all 6 cases BW scores higher than CT, and in 3 cases it reaches 100%. This means that BW is more effective than CT in identifying more severe problems.

Fallout. In all cases BW is better than CT (a smaller fallout means a smaller proportion of errors being produced). The highest fallout for BW is 33%, and in 4 cases fallout is 0.

E–measure. In 3 cases BW is better than CT, and viceversa. This metric ranges from 45 to 86%, and BW tends to stay in the lower part of the range. This is because e–measure is a balanced combination of precision and sensitivity, and the lower sensitivity of BW reflects also here.

But when we consider E_3 then BW scores always better than CT.

Severity values. In all cases the mean severity and the number of problems with highest severity are higher with BW. Therefore BW is more suitable to identify important problems.

5 Conclusions

Based on the sampled data, the barriers walkthrough method appears to be more valid and more useful than conformance tests.

An additional benefit of the method is its role in educating evaluators: after running evaluations with the BW method students become more knowledgeable of accessibility and assistive technology than when doing a dry checklist evaluation using the conformance test.

Additional investigations are needed to determine the reliability of the BW method and to generalize these results to a wider population of evaluators and web sites than those considered in the sample. Consider also that there are disturbance factors: not all the teams analyzed the same set of pages, web sites may have changed across two different evaluations, the cleverness and knowledge of the teams differ, and there might be judging bias. Only a larger study, involving a larger sample of reports, could lead to significant results. Nevertheless even such a small scale experiment provides encouraging results for a more wide spread adoption of the barriers walkthrough method and hopefully contribute to improved accessibility testing practices.

Acknowledgments. I would like to thank Martin Hitz for his help in improving an earlier version of this paper.

References

1. G. Brajnik. Accessibility assessments through heuristic walkthroughs. In *HCItaly 2005 — Simposio su Human–Computer Interaction*, Rome, Italy, Sept. 2005. www.dimi.uniud.it/giorgio/publications.html#hcihw05.
2. Giorgio Brajnik. Web accessibility testing with barriers walkthrough. www.dimi.uniud.it/giorgio/projects/bw, March 2006a. Visited Mar. 2006.
3. Giorgio Brajnik. Simulazione euristica per la verifica dell'accessibilità. www.dimi.uniud.it/wq/metodo-barriere.html, Feb 2006b. Visited Mar. 2006.
4. DRC. Formal investigation report: web accessibility. Disability Rights Commission, www.drc-gb.org/publicationsandreports/report.asp, April 2004. Visited Jan. 2006.
5. European Committee for Standardization. Web accessibility certification workshop. www.cenorm.be/cenorm/businessdomains/businessdomains/isss/about_isss/draft_cwas.asp, March 2006.
6. Italian Government. Requisiti tecnici e i diversi livelli per l'accessibilità agli strumenti informatici. www.pubbliaccesso.it/normative/DM080705.htm, July 2005. G. U. n. 183 8/8/2005.
7. W.D. Gray and M.C. Salzman. Damaged merchandise: a review of experiments that compare usability evaluation methods. *Human–Computer Interaction*, 13(3): 203–261, 1998.
8. S.L. Henry and M. Grossnickle. *Accessibility in the User–Centered Design Process*. Georgia Tech Research Corporation, Atlanta, Georgia, USA, 2004. On–line book, www.UIAccess.com/AccessUCD.
9. M. Hertzum and N.E. Jacobsen. The evaluator effect: a chilling fact about usability evaluation methods. *Int. Journal of Human–Computer Interaction*, 1(4):421–443, 2001.
10. European Information and Communications Technology Industry Association. White paper on eAccessibility. www.eicta.org/press.asp?level2=41&level1=6&level0=1&docid=564, Oct 2005.
11. B. Kelly, D. Sloan, L. Phipps, H. Petrie, and F. Hamilton. Forcing standardization or accomodating diversity? A framework for applying the WCAG in the real world. In S. Harper, Y. Yesilada, and C. Goble, editors, *Int. Cross–Disciplinary Workshop on Web Accessibility, W4A*, pages 46–54, Chiba, Japan, April 2005. ACM.
12. Jakob Nielsen. *Usability engineering*. Academic Press, Boston, MA, 1993.
13. Parlamento Italiano. Disposizioni per favorire l'accesso dei soggetti disabili agli strumenti informatici. www.parlamento.it/parlam/leggi/04004l.htm, Jan. 2004. Legge del 9 gennaio 2004, n. 4.
14. J. Preece, Y. Rogers, and H. Sharp. *Interaction design*. John Wiley and Sons, 2002.
15. John Slatin and Sharron Rush. *Maximum Accessibility: Making Your Web Site More Usable for Everyone*. Addison-Wesley, 2003.
16. W3C/WAI. How people with disabilities use the web. World Wide Web Consortium — Web Accessibility Initiative, w3.org/WAI/EO/Drafts/PWD-Use-Web/20040302.html, March 2004.

Test Case Description Language (TCDL): Test Case Metadata for Conformance Evaluation

Christophe Strobbe[1], Sandor Herramhof[2], Evangelos Vlachogiannis[3], and Carlos A. Velasco[4]

[1] Katholieke Universiteit Leuven, Kasteelpark Arenberg 10, B3001 Heverlee-Leuven, Belgium
christophe.strobbe@esat.kuleuven.be
[2] University of Linz "integriert studieren - integrated study" (i3s3), Altenbergerstrasse 69, A-4040 Linz , Austria
Sandor.Herramhof@jku.at
[3] University of the Aegean, Voulgaraktonou 30, GR11472 Exarchia, Athens, Greece
evlach@aegean.gr
[4] Fraunhofer-Institut für Angewandte Informationstechnik (FIT), Schloss Birlinghoven, D53757 Sankt Augustin, Germany
Carlos.Velasco@fit.fraunhofer.de

Abstract. Automatic benchmarking of evaluation and repair tools (ERT) has been recently the subject of several studies as there is a growing interest because of legal and commercial issues on Web compliance with different criteria and standards. This paper addresses the development of a description language targeted to formally represent test case metadata. This language was used to develop a WCAG 2.0 test suite that will support the benchmarking of ERT with regard to the aforementioned W3C recommendation.

1 Introduction

Since the publication of WCAG 1.0 [4], many evaluation and repair tools (ERT) have been produced to check individual web pages or complete web sites against accessibility guidelines. Each tool has its own features and strengths, but comparing these tools is a complex and time-consuming task. There have been several efforts, for example by Brajnik [1], Melody Ivory [10,11,12] and others. Evaluations of ERT software often rely on comparing evaluation reports of existing web sites. Yvory and Chevalier [11] used sections of existing web sites, asked experienced developers to modify the web pages with the help of one or more tools, and finally asked users to perform certain information-seeking tasks on the original and modified sites. These websites are not available for other researcher to repeat the study with other tools. Brajnik [1] used samples from existing websites which are also unavailable for others to repeat the research. However, Brajnik claims that his method should lead to the same results with different web pages. Nevertheless, these evaluation methods introduce an additional unknown value in an equation where the only unknown value should be the ERT software. This unknown value is the combination of (1) which

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 164 – 171, 2006.

accessibility guidelines are violated, and (2) how often each of these guidelines is violated. Brajnik [1] uses the union of the sets of violations detected by the tools as a basis for evaluating the completeness of each tool. The disadvantage of this approach is that the false negatives produced by the tools distort the evaluation: the assessment of completeness can only be reliable if the number of false negatives that the tools have in common is low. It is therefore necessary to use a suite of test files for which each violation of an accessibility guideline is known and documented. The evaluation report of an ERT product can then be compared with the test suite documentation to determine which guidelines are covered by the product, which guidelines trigger false positives, etcetera. The test suite also makes the evaluation repeatable, because, unlike web sites, a test suite does not change unexpectedly.

The European project BenToWeb[1] has a goal to create test suites for the upcoming Web Content Accessibility Guidelines (WCAG) 2.0. In this context, a test suite is a collection of test cases; each test cases maps to a WCAG 2.0 success criterion and consists of one or more test files and a metadata file. The metadata are encoded in an XML vocabulary specially created for this purpose: Test Case Description Language (TCDL). This Test Case Description Language should not be confused with the TCDL specification submitted to the W3C by IBM in 2003[2]: IBM's TCDL was designed to describe a set of test materials for a formal specification, while BenToWeb's TCDL describes scenarios and other metadata for individual test cases.

2 TCDL: Purpose and Description

TCDL serves two purposes. First, it allows test suite developers to save the metadata that are necessary for testing accessibility evaluation tools. Each test cases maps to a specific "rule", for example a WCAG 2.0 success criterion or WCAG 1.0 checkpoint, and either passes or fails that rule at specific locations in the code. These metadata can then be compared to the output of an accessibility evaluation tool to check if the tool covers the "rules" defined in the test cases, and to check for false positives and false negatives. Second, TCDL supports the definition of test scenarios, when needed, to validate the test cases during the development process. Each test case needs to be reviewed for obvious quality assurance reasons, but it is also possible to define scenarios for end-user testing.

TCDL is defined in W3C XML Schema[3] and uses the namespace http://bentoweb.org/refs/TCDL1.1; the XML Schema is available at http://bentoweb.org/refs/schemas/tcdl1.1.xsd and contains detailed documentation.

Each TCDL file has a unique identifier and consists of five major sections: formal metadata, "technologies", the "test case" (test files and test scenarios), "rules" (data related to the relevant guidelines), and namespace mappings.

---

[1] `http://www.bentoweb.org/`

[2] `http://www.w3.org/TR/2005/NOTE-test-metadata-20050914/`

[3] XML Schema is defined at `http://www.w3.org/TR/2004/REC-xmlschema-1-20041028/` and `http://www.w3.org/TR/2004/REC-xmlschema-2-20041028/`.

2.1 Formal Metadata

Formal metadata are mostly “administrative” information (as opposed to semantic metadata). They consist of a title, a description, creation date and author, rights (usually copyright), status information, and data about sources from which materials were borrowed. Many of these data are optional.

The title contains merely a short title that allows a quick identification of the test case; it is not a unique identifier for machine processing. The description is a summary of the test files and how they are to be used in the test case. It describes to the evaluator or the developer of test scenarios what to expect, and, especially when the test files require user interaction, for example through forms, what will happen when interacting with the file. It can be very simple, for example: “A document with an `img` element that displays a cat. The `img` element does not have an `alt` attribute.” It can also be more complex, for example when describing form interaction: “A page with a form that has two `fieldsets`. The first `fieldset` contains a group of radio buttons, and the user is required to make a choice (there is no default). Instructions above the form explain that required fields have labels in red. The red text for the first `fieldset` is created by means of CSS and does not include a textual or character cue to tell the user that the field is required, so it is very difficult for users of current screen readers to find out which form fields are required. The form relies on the user's ability to recognize red text.”

For language, creator and rights related to the test materials in the test case, TCDL reuses Dublin Core metadata elements (dc:lang, dc:creator and dc:rights, respectively).

The status of the test case can be tracked as the test case moves from “draft” through evaluation to “accepted”. Each test case starts out as a draft and is then reviewed by another accessibility or HCI expert. If any issues are found, the status changes to “pending bugfix” and the test case is sent back to the test case author. The test case author fixes the issue and sets the status back to “draft”. It is also possible that the test case contains scenarios for end-user evaluation (explained below); when these scenarios are finalised, the status is set to “accepted for end user evaluation”. A test case evaluation framework can select these test cases and present them to end users. (BenToWeb’s test case evaluation framework is described by Herramhof et al [7].) After evaluation and when all data are definitive, the status is set to “accepted”.

If test materials are borrowed — for example from existing test suites — the source, rights, and any other comments can be recorded in an optional section called “source”.

Finally, there is also an optional “version” field, which is not used in BenToWeb, but which may of value to other users of TCDL.

2.2 Technologies

TCDL uses the term “technology” in the same sense as WCAG 2.0: “a markup language, programming language, style sheet, data format, or API”[4]. Each test case implements a specific technology feature, for example an HTML element such as `img`,

[4] `http://www.w3.org/WAI/GL/WCAG20/WD-WCAG20-20060407/appendixA.html#technologydef`

that is relevant to the "rule" for which the test case is defined. TCDL provides the ability to store the following information about the technologies in a test case:

- the formal specification or recommendation that defines the technology (by name and URL; mandatory);
- the technology features (for example, which XHTML elements and/or attributes) that are relevant to the test (optional);
- a reference (URL) to a section in the formal specification, or a quote from the formal specification that describes the technology features that are relevant to the test (optional).

2.3 Test Case

The "test case" section consists of four subsections: purpose, required tests, test files and preconditions. The preconditions are an optional section.

Purpose. The "purpose" is a mandatory section that provides a brief explanation why the test was developed. It explains the expected evaluation result ("pass" or "fail") in regard to the relevant "rule" (WCAG checkpoint or success criterion). It is not meant to repeat the "rule", but in complex test cases, it can state that certain aspects of the test file(s) are not meant to be tested. For example, in a test of text alternatives for image map areas, the purpose can state that only the alternative text should be evaluated, and not the colour contrast of the image itself.

Required Tests. This section describes what type of evaluation is necessary to move the test case from "draft" to "accepted". A test case can be evaluated in one or more ways: by end users, one accessibility expert, a group of accessibility experts, or automated tools. These options are called "test modes" in TCDL; the definition of one or more test modes is mandatory.

Depending on the test mode, it is also possible to define test scenarios. Defining test scenarios is optional, except if the test mode is end-user evaluation. A test scenario consists of the following parts:

- user guidance (for example, advice on features of a user agent that should be supported by and enabled in a browser);
- a question that the user is expected to answer (yes/no question, open question, Likert scale, multiple choice, or a yes/no/other question);
- the experience that the user should have with certain user agents or assistive technologies to evaluate the test, and;
- optionally, disabilities to which the test is relevant.

The third part, experience, defines the types of user agents, assistive technologies and devices with which the test files should be evaluated. User agents are usually browsers and browser plug-ins; devices are usually PCs, but PDAs and cell phones could also be relevant. For each of these softwares or platforms, it is possible to define the minimal experience level that a user should have with them. A test case evaluation framework can then match these software, platforms and experience levels to user profile data and present relevant scenarios to the end user.

For each category (user agent, assistive technology, device) it is possible to specify only the "type" (for example, browser, screen reader, magnification software, speech recognition software), but it is also possible to specify the products and (minimal) product version with which the test case should be evaluated. If several products of the same "type" are specified (for example, the browsers Internet Explorer 6.0, Firefox 1.0 and Opera 7.54), these are in an OR relationship. The different categories, however, have an AND relationship. For example, if a test case requires a browser (for example, one of Internet Explorer 6.0, or Firefox 1.5) and a screen reader (for example, one of JAWS 7.0 or Window-Eyes 5.5), the test case will be presented to users with both Internet Explorer 6.0 and JAWS 7.0, or with both Internet Explorer 6.0 and Window-Eyes 5.5, or with both Firefox 1.5 and Window-Eyes 5.5, etcetera, but not to users who have a older version of these products, or to users who have SuperNova (as a screen reader), or Safari (as a browser) or to users who don't have a screen reader. If the scenario also specifies disabilities; the test case evaluation framework can use this information as additional criteria when matching test scenarios to user profiles.

When a user is presented with a test that matches their setup, it may be necessary to provide extra guidance about certain features that need to be enabled for the evaluation to be useful. Many browsers and assistive technologies can be customised, and some tests may rely on features that the user happens to have disabled. For example, not all users have JAWS set up so that it reads the expansions of acronyms.

Most importantly, the user is asked to interact with the test file (or files) and to answer a question. TCDL supports several types of questions: simple yes/no questions, Likert scales, open-ended questions, yes/no questions with a third option for the user to fill in, and multiple-choice questions. The test case evaluation framework transforms this question into an HTML form; the answer is stored in a database or other persistence mechanism for later analysis.

Test Files. A test case has one or more test files. For most "rules" or WCAG 2.0 success criteria, one file is sufficient. However, certain success criteria only apply to sets of files, for example, success criteria about consistency of navigational mechanisms or about information about a user's location in a website. Usually, it is sufficient to provide only the URL of the test file. In some cases, however, it is useful to test server responses that result from certain specific values in HTTP request headers or the HTTP request body. For this reason, TCDL also supports the definition of specific mediatypes, enctypes, HTTP request headers and HTTP request body. If more than one test file is specified, it is also possible to define if the sequence in which the files are navigated is meaningful. Supporting files such as included images, external style sheets and dummy pages to which forms are submitted (if only the form is relevant to the test) etcetera, are not documented.

Previous research [9] shows that test cases with multiple files appear to be something new in test suites for Web standards.

Preconditions. Preconditions are "conditions that must be met before the test can be successfully executed" [6]. This kind of conditions is used in the current HTML test suite for WCAG 2.0[5], where they are called "prerequisite tests". BenToWeb does not define preconditions because the BenToWeb test suite relies on a different mapping

---

[5] `http://www.w3.org/WAI/GL/WCAG20/tests/`

mechanism than the current HTML test suite for WCAG 2.0: test cases map directly to WCAG 2.0 success criteria, instead of through the intermediate step of techniques.

2.4 Rules

TCDL uses “rules” as a generic term for WCAG 1.0 checkpoints, WCAG 2.0 success criteria, BITV provisions, and similar accessibility requirements. The “rules” section defines the mapping between the test case and specific accessibility requirements. The rules do not provide a direct reference to the relevant online source because it may not always be possible to reference “rules” directly (for example, because they are not available in HTML) or because they might move to a different URL. The rules reference a rule ID in a “Rulesets XML” file (see Rulesets XML RDDL document[6]), which in turn references the relevant rule in an online source.

For each rule listed in this section, the following information can be provided:

- whether the rule is the “primary” rule of the test case or only listed for informative purposes (for example, some test cases uses invalid markup in tests that are not about validity, so it is useful to document that checkpoints about validity should not be applied to the test case) (optional);
- the identifier of the rule in “Rulesets XML” (mandatory);
- the locations in the test file that are relevant to the rule, with URL, line number, column number and/or XPath expression [5] (optional);
- whether the test case passes or fails the success criterion or rule (mandatory);
- the functional outcome of the test: a description of why the test case passes or fails, in terms that relate to a user's experience (as opposed to technical comments about source code) (mandatory);
- technical comments on the test: a technical description of why the test case passes or fails, and other technical information (for example, issues in certain browsers or assistive technologies).

2.5 Namespace Mappings

Finally, there is an optional section for “namespace mappings”, which is important for test cases based on test files that use XML-based technologies. The technology features (see the section on Technologies above) used in a test case exist in an XML namespace [2] and should be referenceable with an XPath expresssion. The normal mechanism for mapping namespaces to URIs by declaring them with `xmlns:prefix="namespaceURI"` cannot be used for this, because it is necessary to be able to define empty prefixes and to extract the namespace mappings with a common XML API.

2.6 Differences with IBM's TCDL Submission to W3C QA

BenToWeb's TCDL was developed completely independently from the TCDL proposal submitted by IBM [8] to the W3C Quality Assurance Working Group (QA WG) in 2003. The goal of IBM's TCDL is to catalogue most of the test materials that a

---

[6] `http://bentoweb.org/refs/rulesets`

W3C Working Group would provide. It is intended to be used in a test lab "to set up and run the test materials against one or more test subjects" and "supports automated setup of the system(s) used for testing, automated running of test cases, automated comparison of results, and automated cleanup of the system(s)". Like BenToWeb's TCDL, IBM's TCDL is an XML vocabulary, so metadata can be transformed into other formats for human or machine consumption.

An IBM TCDL catalogue contains a list a test cases, whereas a BenToWeb TCDL contains metadata for an individual test case only. BenToWeb's TCDL is also different because each test case references at least two specifications (WCAG 2.0 or another ruleset, and a specification of a technology such as XHTML or CSS), and one of these (WCAG 2.0) cannot be tested fully automatically.

Unfortunately, it is not possible to make a detailed comparison of the two languages because the XML Schema for IBM's TCDL submission is not available.

3 Conclusions and Future Work

Even though the BenToWeb consortium is not the first organisation that develops accessibility test suites, the Test Case Description Language appears to have no predecessors that fulfil the same requirements. TCDL has undergone almost a year's implementation experience through the creation of test case descriptions, a desktop TCDL editor, transformations into XHTML, and implementation in the test case evaluation framework (see Herramhof et al [7]). TCDL implements the test metadata defined by the W3C's Quality Assurance Working Group [6] in order to support use cases outside BenToWeb. Through informal contacts, the W3C's ERT Working Group has expressed interest in this language, and there are plans to submit it to the W3C.

Acknowledgements

This work has been undertaken in the framework of the project BenToWeb — IST-2-004275-STP — funded by the IST Programme of the European Commission.

References

1. Brajnik, G (2004). Comparing accessibility evaluation tools: a method for tool effectiveness. Univ Access Inf Soc 3: pp. 252-263. DOI: 10.1007/s10209-004-0105-y
2. Bray, T, Hollander, D, Layman, A (1999). Namespaces in XML — World Wide Web Consortium 14-January-1999. Available at: `http://www.w3.org/TR/REC-xml-names/`
3. Caldwell, B, Chisholm, W, Slatin, J, Vanderheiden, G, White, J (eds) (2005). Web Content Accessibility Guidelines 2.0, W3C Working Draft 30 June 2005. World Wide Web Consortium. Available at: `http://www.w3.org/TR/WCAG20/`
4. Chisholm, W, Vanderheiden, G, and Jacobs, I (eds) (1999). Web Content Accessibility Guidelines 1.0, W3C Recommendation 5-May-1999. World Wide Web Consortium. Available at: `http://www.w3.org/TR/WCAG10/`

5. Clark, J, DeRose, S (1999). XML Path Language (XPath) 1.0 — W3C Recommendation 16 November 1999. Available at: http://www.w3.org/TR/xpath
6. Curran, P, Dubost, K (2005). Test Metadata — W3C Working Group Note 14 September 2005. Available at: http://www.w3.org/TR/2005/NOTE-test-metadata-20050914/
7. Herramhof, S, Petrie, H, Strobbe, S, Vlachogiannis, E, Weimann, K, Weber, G, Velasco C A (2006). Test Case Management Tools for Accessibility Testing. Computers Helping People with Special Needs. 10th International Conference, ICCHP 2006, Linz, Austria, 12-14 July 2006, Proceedings
8. Marston, D (2003). Test Case Description Language 1.0 — Submisson to QA Working Group 12 October 2003. Available at: http://www.w3.org/QA/WG/2003/10/tcdl-20031012.html
9. Strobbe, C (2005). Test-suites' State of the Art and Quality Assurance Methods for W3C Recommendations. BenToWeb deliverable D 4.1. Available at: http://www.bentoweb.org/html/BenToWeb_D4.1.html
10. Ivory, M Y, Sinh, R R, and Hearst, M A (2001). Empirically validated web page design metrics. Proceedings of the Conference on Human Factors in Computing Systems (Seattle, WA, March), pp. 53–60. New York, NY: ACM Press
11. Ivory, M Y, and Chevalier, A (2002). A Study of Automated Web Site Evaluation Tools. Technical Report UW-CSE-02-10-01. University of Washington, Department of Computer Science and Engineering. http://ubit.ischool.washington.edu/pubs/tr02/toolstudy.pdf
12. Ivory-Ndiaye, M Y (2003). An Empirical Approach to Automated Web Site Evaluation. Journal of Digital Information Management, 1(2), June 2003, pp. 75–102.

The BenToWeb XHTML 1.0 Test Suite for the Web Content Accessibility Guidelines 2.0

Christophe Strobbe[1], Sandor Herramhof[2], Evangelos Vlachogiannis[3], Johannes Koch[4], and Carlos A. Velasco[4]

[1] Katholieke Universiteit Leuven, Kasteelpark Arenberg 10, B3001 Heverlee-Leuven, Belgium
christophe.strobbe@esat.kuleuven.be
[2] University of Linz "integriert studieren - integrated study" (i3s3), Altenbergerstrasse 69, A-4040 Linz, Austria
sandor.herramhof@jku.at
[3] University of the Aegean, Voulgaraktonou 30, GR11472 Exarchia, Athens, Greece
evlach@aegean.gr
[4] Fraunhofer-Institut für Angewandte Informationstechnik (FIT), Schloss Birlinghoven, D53757 Sankt Augustin, Germany
{Johannes.Koch, Carlos.Velasco}@fit.fraunhofer.de

Abstract. This paper presents a detailed description of the work carried out under the umbrella of the EU-funded project BenToWeb to develop a complete XHTML 1.0 Test Suite in regard to conformance with the Web Content Accessibility Guidelines 2.0 from the W3C. This initial work covered the Working Draft version of June 2005. A thorough evaluation involving end users is carried out at the moment of writing this paper.

1 Introduction

The EC-funded project BenToWeb (Benchmarking Tools and Methods for the Web[1]) has a goal to develop test suites for the forthcoming Web Content Accessibility Guidelines 2.0. The development of these test suites serves a dual purpose: first, they support the WAI Working Groups in the development of support documents, such as technology-specific techniques, for WCAG 2.0; second, the test suite can be used to benchmark accessibility evaluation and repair tools (ERT). WCAG 1.0 has general techniques (core techniques), HTML techniques and CSS techniques; WCAG 2.0 will probably have general techniques, HTML techniques, CSS techniques, client-side scripting techniques and server-side techniques. The BenToWeb test suites will not cover all of these technologies, but at least XHTML 1.0 + CSS 2.0.

2 Structure of the Test Suite

In this context, a test suite is not a set of tests than can be used to validate web content, but a set of test files with accompanying metadata both for human and machine

[1] http://www.bentoweb.org/

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 172 – 175, 2006.

consumption. The test suite is a collection of “test cases”, where a test case consists of one or more XHTML files that implement or fail a requirement specified by a WCAG 2.0 success criterion, and an accompanying metadata file. The metadata are recorded in an XML format specially created for this purpose: Test Case Description Language (TCDL) [3]. The metadata include a description of the test file or files, the purpose of the test, a link to the success criterion that the test case is meant to implement or fail, a statement on whether the test case passes or fails the success criterion, and sometimes also test scenarios for the purpose of validation (especially for end-user evaluation).

For each WCAG 2.0 success criterion, at least two test cases need to be created: at least one that fails and at least one that passes the success criterion. When the test suite is complete and validated, running the test files through an accessibility evaluation tool should then provide data on the completeness of the tool’s coverage of WCAG 2.0 and whether it generates false positives and negatives.

There can be several types of test cases. “Atomic test cases” address only one success criterion and use only a single XHTML file (supporting files such as images or CSS style sheets do not count in this context). However, some accessibility requirements apply to sets of files instead of single files: WCAG 2.0 contains success criteria about consistency of navigational elements, and about information on a user’s location in a web site. Test cases for these success criteria use multiple XHTML files and are called “compound test cases”. However, it is also possible to create “complex test cases”: test cases that implement or fail multiple success criteria. At the time of writing, the BenToWeb test suite contains no test cases that are identified as “complex”, but the Test Case Description Language supports this. Previous research has shown that the concept of “complex test cases” is new in the area of accessibility test suites [2].

3 Development of the Test Suite

The development process requires that each test case moves through several steps before it is finally accepted in the test suite. Each test case starts out as a draft and is then reviewed by another accessibility or HCI expert. If any issues are found, the test case is sent back to the test case author. It is also possible that the test case contains scenarios for end-user evaluation. A test case evaluation framework can select these test cases and present them to end users. (BenToWeb’s test case evaluation framework is discussed elsewhere [1].) After evaluation and when all data are definitive, the test case is finally “accepted” into the test suite. The Test Case Description Language contains metadata related to the status of a test case.

At the time when development of the test suite started, the Web Content Accessibility Guidelines Working Group (WCAG WG) had only defined a relatively small set of HTML techniques for WCAG 2.0. At that time, there was also an HTML Test Suite[2] which defined tests and contained examples of files that passed or failed these tests. Since the techniques and the tests were still at an early stage, BenToWeb chose to map the test cases directly to success criteria instead of the existing techniques and

[2] `http://www.w3.org/WAI/GL/WCAG20/tests/`. The test suite still exists at the time of writing, but the WCAG WG is integrating tests into the techniques documents. The ERT and WCAG WGs might set up a joint task force for the development of tests for WCAG 2.0.

tests. Test case authors were free to draw on any documentation of techniques or failures they could find, regardless whether the source was WCAG or not.

4 Current State of the Test Suite

At the time of writing, the test suite contains 481 test cases for the 67 success criteria in the 30 June Working Draft of WCAG 2.0, the current draft when the development of the test suite started. These test cases contain over 530 XHTML test files (or JSP files that generate XHTML), which often uses supporting files, such as JavaScript, CSS, GIF, JPEG, WMV (audio/video), WMA, WAV, MP3 and Java applets.

Some success criteria have only two test cases, while others have more than thirty. The variability in the number of test cases per success criterion is often related to the number of XHTML elements or attributes that can be used to pass or fail a success criterion: for example, this number is much higher for Guideline 1.3 Level 1 Success Criterion ("Structures within the content can be programmatically determined") than for Guideline 3.1 Level 3 Success Criterion 3 ("A mechanism for finding the expanded form of acronyms and abbreviations is available"). Another reason is that some success criteria in the June 2005 draft of WCAG 2.0 were unfinished or open to interpretation. Unfinished success criteria included some success criteria for Guideline 1.4 (contrast between foreground and background), success criteria for Guideline 2.3 (flashing content), and the Guideline 4.2 success criterion that depends on Guideline 2.3. Success criteria that were open to interpretation included several success criteria for Guideline 3.2, where the phrases "programmatically determined", "change of context" and "initiated only by user action" were either too strict or too loose for what was intended. These two types of ambiguity were handled in the Test Case Description Language by setting the expected evaluation result of the test files to "cannot tell" (instead of "pass" or "fail"). Some issues, especially about changes of context, were fed back to the WCAG Working Group through their mailing list. Some success criteria are not covered by the test suite, because they are not applicable to XHTML, because they depend on the definition of a baseline (which BenToWeb chose not to do), or even because no technique to implement the success criterion could be found. On the other hand, there are a few test cases for success criteria that existed in previous drafts of WCAG 2.0, but which were removed from the June 2005 draft: the much contested success criteria about well-formedness (for XML) and validity in Guideline 4.1. For XHTML, checking well-formedness and validity against a DTD is very straightforward, so test cases that check whether an accessibility evaluation tool supports validation are useful components in a test suite.

5 Future Work

At the time of writing, the WCAG Working Group has published a new draft of WCAG 2.0 (November 2005) and is preparing a Last Call Working Draft. The most important implication of this for the test suite — and one which as inherent in the mapping of success criteria to a working draft — is that the mapping of the test cases to success criteria will need to be updated to a stable version of WCAG 2.0 (Last Call

Working Draft or, eventually, Recommendation). It also means that test cases will need to be created for success criteria that were added since the Working Draft of 30 June 2005. Moreover, some existing test cases will need to be reviewed because the related success criteria have changed. Other test cases will disappear as the related success criteria no longer exist.

Another important task is the creation of more test cases for all success criteria, in order to create a larger set from which subsets can be generated. This is necessary to avoid that developers of ERT “optimize” their software for the test suite, in order that they score high in benchmarks without matching improvements in their evaluation of real web sites. Creating a bigger set of test cases from which test suites can be generated would ensure that the actual test set used to benchmark an evaluation tool is not predictable.

Acknowledgements

This work has been undertaken in the framework of the project BenToWeb — IST-2-004275-STP — funded by the IST Programme of the European Commission.

References

1. Herramhof, S, Petrie, H, Strobbe, S, Vlachogiannis, E, Weimann, K, Weber, G, Velasco C A (2006). Test Case Management Tools for Accessibility Testing. Computers Helping People with Special Needs. 10th International Conference, ICCHP 2006, Linz, Austria, 12-14 July 2006, Proceedings
2. Strobbe, C (2005). Test-suites' State of the Art and Quality Assurance Methods for W3C Recommendations. BenToWeb deliverable D 4.1. Available at: `http://www.bentoweb.org/html/BenToWeb_D4.1.html`
3. Strobbe, C, Herramhof, S, Vlachogiannis, E, Velasco C A (2006). Test Case Description Language (TCDL): Test Case Metadata for Conformance Evaluation. Computers Helping People with Special Needs. 10th International Conference, ICCHP 2006, Linz, Austria, 12-14 July 2006, Proceedings

An Environment for Defining and Handling Guidelines for the Web

Barbara Leporini, Fabio Paternò, and Antonio Scorcia

ISTI – CNR
56124 Pisa, Italy
{barbara.leporini, fabio.paterno, antonio.scorcia}@isti.cnr.it

Abstract. Several accessibility and usability guidelines are proposed more and more. In this paper we present an environment for defining, handling and checking guidelines for the Web. The goal of such a tool is to support developers and evaluators in flexibly handling multiple sets of guidelines, which can be dynamically considered in the evaluation process. In particular, an interactive editor has been designed to assist the evaluators in abstracting and specifying new and existing guidelines in our XML-based Guideline Abstraction Language (GAL), which are then stored in external files. Our tool is able to check any guidelines specified in this language without requiring changes in its implementation.

1 Introduction

Design guidelines are increasingly proposed in order to improve and make consistent user interfaces, in particular for better supporting accessibility and usability. Indeed, several guidelines have been proposed in literature for general user interfaces (e.g., [6,7 and 8]) as well as for Web interfaces (e.g. [5,9 and 10]). Thus, in order to assure a certain level of accessibility and usability, developers have to orientate themselves among several sets of guidelines. To this end, we have developed a tool, MAGENTA (Multi-Analysis of guidelines by an Enhanced Tool for Accessibility), which supports various sets of guidelines. One of its mean features is to be independent of the guidelines specification. This means that if the guidelines are expressed by an abstract language that we propose and are stored in external files then the tool can check any of them without modifications in its implementation. In this way, it is possible to:

- Specify any kind of guideline based on X/HTML tags or CSS attributes;
- Change any existing guideline simply by handling the corresponding XML-based specification;
- Group and separate several guidelines into various sets.

In our first experiences with the tool, we soon realised that if files are handled manually by developers or evaluators then many problems could occur:

- File editing should be made manually by an expert and it can be rather tedious;
- Editing could introduce some syntactical errors, unless the XML file is not parsed by an XML validator;

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 176 – 183, 2006.

- Adding new guidelines as well as changing existing ones could require a lot of time and effort;
- The experts who handle the XML file and guidelines should know tags, attributes and properties very well.

For all these reasons, we have extended our tool with a graphical environment thought for assisting developers and evaluators in proposing, specifying and modifying guidelines. In this paper we present an environment, which includes a Guideline Editor (GE) developed for assisting developers in defining new guidelines and in handling those already existing (such as those for accessibility and usability for the Web). In particular, such an editor has been integrated into the tool MAGENTA in order to improve its flexibility.

The paper is structured as follows: firstly we introduce a short description of the tool MAGENTA; next, the Guideline Abstraction Language (GAL) is described. In section 4, an interactive editor for defining guidelines according to the GAL is proposed and discussed.

2 The MAGENTA Tool

The MAGENTA tool has been developed with the intent to check whether a Web site is accessible and usable and provide support to improve it. To this end, the tool checks how satisfactorily the selected guidelines are applied to the indicated Web pages. This is obtained through automatic identification of the checkpoints associated with each guideline and analysis of the associated constructs and attributes to check whether they provide the necessary information. Currently, MAGENTA supports three sets of guidelines for the Web: a set of guidelines for visually-impaired users [3], WCAG 1.0 [9], and the guidelines associated with the Italian accessibility law [1]. Firstly, the guideline set to be considered can be selected. Then, through a list of checkboxes, the evaluator can decide which guidelines of the selected set(s) have to be checked.

The tool is not limited to checking whether the guidelines are supported but, in case of failure, it also provides support for modifying the implementation in order to make the resulting Web site more usable and accessible. Thus, when an error is detected, the tool points out what parts of the code are problematic and provides support for corrections indicating what elements have to be modified or added. The process is not completely automatic because in some cases the tool requires designers to provide some information that cannot be generated automatically. The tool indicates the parts that must be corrected (i.e. tags or attributes) and developers write those parts or values that can not be added automatically (e.g. specification of proper link contents or proper name for frames).

MAGENTA has been developed considering the limitations of most current tools, in which the guidelines supported are specified in the tool implementation, with all the associated issues mentioned in the introduction (e.g., having to change the tool implementation each time a guideline is modified or added). In our work we aim to provide a tool (MAGENTA) independent of the guidelines to check. Practically, the tool is able to deal with guidelines which are defined externally from the tool. The guidelines are specified in an XML file in a specific abstract language whose basic constructs can be interpreted by our tool at run-time. In this way, the tool is able to adapt the checking procedure to various guideline sets as well as new set of guidelines, which are specified by the evaluators and it does not require to be modified and recompiled. This opens up many interesting opportunities to developers and evaluators who can indicate flexibly what guidelines they want to check even if they do not belong to those already existing and available. However, as shortly described in the introduction, handling manually XML-based guideline specification files could generate several difficulties and errors. Thus, we decided to extend the guideline definition support by adding a new module: the Guideline Editor, which has been added to the MAGENTA functionalities in order to improve its flexibility.

3 The Guideline Abstraction Language

Our solution is based on the definition of a language for specifying guidelines that are stored externally to the tool. In general, a guideline, is defined as a rule or principle that provides guidance to appropriate behaviour. Generally speaking, a guideline is expressed in natural language, which automatic tools cannot handle. Hence, in order to support developers in using and customizing various sets of guidelines, our approach is aimed at expressing guidelines so that they can be dealt with automatically. In practice, the solution consists of abstracting guideline statements in a well-defined language. For example, a statement such as “all images in a Web page must have an alternative description” or “Tables should have a short summary description” have to be formalised in some manner that can be handled by the automatic tool. Thus, we require that the guideline statement be abstracted and appropriately specified. For the guidelines specification we use an XML-based language. This implies that the guideline is previously structured and then specified in our XML-based language.

In defining a guideline abstraction, the main features and elements characterizing the guideline itself must be defined. Hence, general features, objects involved in the checking process, and conditions referred to objects must be specified through the XML-based Guideline Abstraction Language (GAL) that we propose. For example, a guideline such as “All links must have clear and significant content” could be abstracted as follows:

```
<a> with href, XHTML, (content not_belong
not_proper_list_values)
```

and

```
<a> with href, XHTML, (for <img> content: Exist alt &
(alt not_belong not_proper_list_values | alt+content
not_belong not_proper_list_values))
```

In this case, checking involves both textual and graphical links, i.e. the criterion has two checkpoints: (1) textual and (2) graphical links. In the former case the content should not belong to a list of non appropriate terms, in the latter case the alt attribute should exist with values not belonging to inappropriate terms.

In brief, each guideline expresses a principle to be complied with by applying one or more conditions at checkpoints. All this information and these conditions have to be structured by a specific guideline abstraction. More details can be found in [4].

Summarising, we define a guideline in terms of a number of elements: checkpoints expressed with objects, operations (such as Exist, Count, Check, Execute) and conditions to be verified. In specific, an XML-Schema [11] has been defined in order to define the general structure of our Guideline Abstraction Language (GAL). The next table reports an excerpt of an example of a guideline expressed with XML tags according to the proposed language for abstracting guideline definition.

Table 1. Excerpt of Example Guideline Specification

```
<name> Italian accessibility law</name>
...
<criterion type="accessibility" target="page">
<description>Links should have an appropriate content
</description>
<checkpoints rel="and">
<cp id="19.1" summary="Textual links" priority="1">
   <cp_descript>...</cp_descript>
   <eval_object code="html" mandatory="yes">
      <object type="tag">a</object>
      <select type="attrib">href</select>
   </eval_object>
   <conditions rel="and">
   <evaluate operator="check" cond="not_belong" iderr="19.1.1">
      <el type="tag">a</el>
      <el type="file">"nolinks.dic"</el>
   </evaluate></conditions>
   </cp>
<cp id="19.2" summary="Graphical links" priority="1">
   <cp_descript>...</cp_descript>
   <eval_object code="html" mandatory="no">
      <object type="tag">a</object>
      <select type="attrib">href</select>
      <select type="tag">img</select>
   </eval_object>
<conditions rel="and">
...
</guideline></gdl_set>
```

4 The Guideline Editor

When a guideline must be defined so that a tool can deal with it automatically, the problem is to transform its description in natural language into technical terms. To this end, we have defined a XML abstract language for describing guidelines. The abstraction process is not easy and requires some effort when implementation languages for the Web are considered. Furthermore, according to the proposed GAL, to define a guideline several XML tags and attributes must be specified. In order to facilitate this process, a graphical editor has been designed and added to the MAGENTA tool, thus enabling even people non particularly expert in handling languages such as X/HTML and CSS to specify the desired guidelines.

The Guideline Editor (GE) has been designed for assisting developers in handling single as well as groups of guidelines. The tool supports new guideline definition and various types of editing. Summarising, the editor offers various useful features in order to facilitate multiple guideline sets management and general utilities. In particular, the guideline editor provides support for:

- Defining or modifying single guidelines;
- Importing guidelines from various sets in order to create customised groups;
- Organising, classifying and browsing guidelines into groups;
- Handling several sets of guidelines simultaneously;
- Writing and repairing external dictionaries used with some guidelines (e.g. proper link content, proper name for frames, specific sections, and so on).

Indeed, the last feature has been thought for some kind of guidelines that need some support in identifying appropriate or not appropriate terms. For instance, if a guideline states "frames should have the attributes 'title' appropriate", then some not appropriate values could be stored in an external file called "dictionary". Such dictionary is used during the evaluation process in order to check that frames have a title attribute which does not belong to the non-appropriate term list. In practise, we verify that the value is not inappropriate (negative control). Other guidelines can be checked with a similar approach.

Regarding manipulation of single guidelines, the tool allows developers to add new ones or modify/delete existing ones. In practice, the graphical editor guides in specifying all the XML tags necessary for defining a guideline (see an example in Table.1). Furthermore, it is also possible to add new guidelines by importing them from other guideline set(s). This allows designers to create customised set(s) of guidelines by reusing already specified ones. More precisely, through the GE for a guideline it is possible to:

- Insert general information on the guideline (short name, description, etc.);
- Select the object to be considered in the guideline (i.e., language, tags and restriction on the object itself).
- Indicate the number of checkpoints needed and their relationship (and | or);
- Insert general information for each checkpoint, such as short name, description and the number of the conditions composing the checkpoint;
- Define the conditions to be verified, by selecting operators and conditions to be applied at a runtime (i.e. when the tool will carry out the evaluation).

The tool user interface has been designed in order to be easy for any user, including those who are not particularly expert with XML and X/HTML languages. In fact, since it is not possible to specify a guideline precisely in a natural language, then the XML elements involved in the guideline must be edited. To this end, when selecting a specific object (tag, attribute or property) a short description is given to the developers. In practise, the tool guides the user in editing all the elements necessary to define a guideline in a technical format. Thus, the tool assists developers in abstracting a guideline and then generates its specification into the XML external file.

The Guideline Editor UI is mainly organized in various parts shown in cascade. Three main functionalities are available:

- *Main Guideline Management*, where all groups of guidelines can be handled (see Figure 1, top); it is possible to create a new set, modify or delete an existing one. The available guideline sets are listed; through the buttons "New", "Edit", "Show" and "Import" all the corresponding activities can be performed.

- *Guideline set*, which is the environment in which a new specific group of guidelines can be defined. General information of the set as well as the list of guidelines already belonging to the group are presented (see Figure 1, bottom).
- *Guideline specification*, where developers can structure and define a single guideline. Firstly, general data such as the short name and description of the current guideline as well as the main object – tag, attribute or property - involved by the guideline can be written and modified; next, all checkpoints and their conditions to be applied or evaluated can be edited through a direct manipulation graphical editor (see Figure 2).

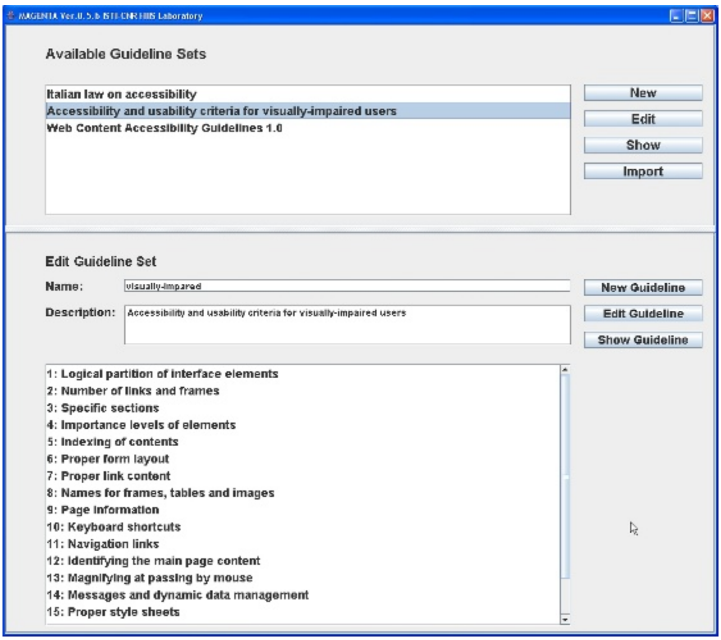


Fig. 1. The interface supporting the creation of a new set of guidelines to check

Some elements can be selected from a list or typed by a combo-box. In practice, the elements to be selected are listed in order to make them available for the evaluator and to avoid typing errors. In this way, the user is guided during the abstraction and definition process. For instance, when an object to be checked has been selected, in the list-box for selecting the attributes related to the element, just its attributes are available. For example, if the element chosen is the tag <a>, all the attributes available in the other list-box are only those related to the tag <a> (e.g., href, tabindex, accesskey, etc.). If no restriction must be applied the "Anyone" item can be chosen.

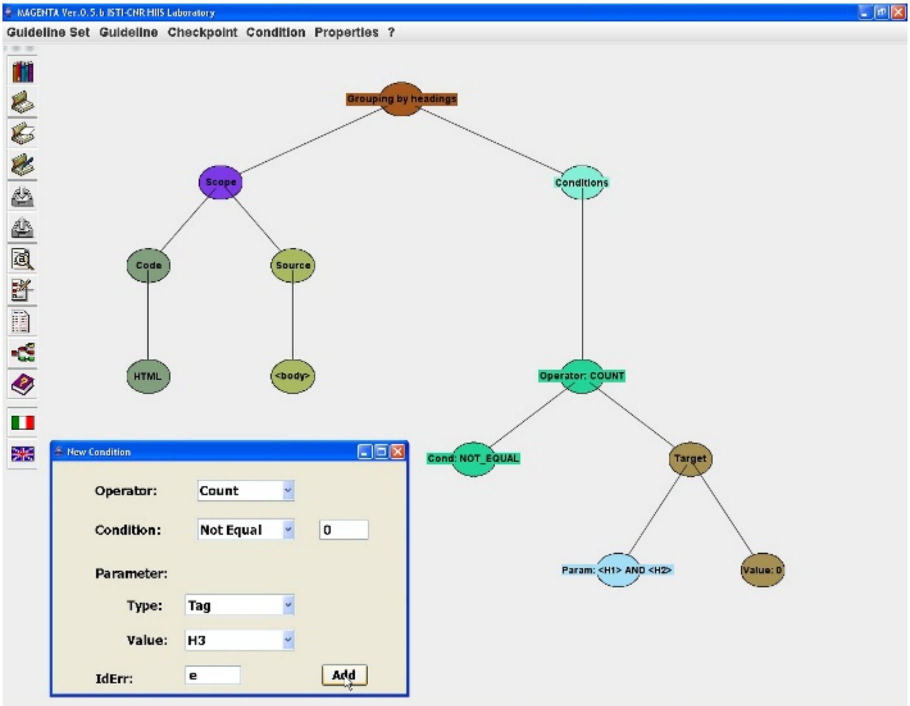


Fig. 2. The interface supporting the definition of a new guideline

When the user has selected all the elements which express the objects to be checked, including its properties, checkpoints and conditions, the editor saves them in the XML-based file associated to the current guideline set by inserting all the elements in the correct tags defined for each guideline.

5 Conclusions

The increasing number of design guidelines proposed for the Web, in particular for accessibility evaluation, makes the implementation of automatic tool working with guidelines more and more complex. In order to develop a tool independent of guideline definition, guidelines should be specified separately and interpreted at runtime. To this end, we propose a Guideline Abstraction Language to define guidelines using an XML structure. These XML-based files can be managed by our MAGENTA tool, which loads and check the guidelines indicated by the developer without requiring modifications to its implementation. Thus, designers can dynamically select and edit the guidelines to check in their Web site. Since abstracting and defining guidelines requires some effort, a graphical Guideline Editor has been integrated in our environment for assisting evaluators in these activities. The user interfaces has been designed to guide developers as much as possible in order to facilitate their task. Currently, the editor supports guidelines for documents implemented in X/HTML and CSS languages, future work is planned to support other Web languages such as SMIL, XForm, and so on.

References

1. Italian Accessibility Law, 2004, Law n. 4, January 9, 2004: Provisions to support the access to information technologies for the disabled
2. http://www.pubbliaccesso.it/normative/law_20040109_n4.htm_20040109_n4.htm
3. Leporini, B., Paternò, F. (2004). Increasing Usability when Interacting through Screen Readers, International Journal Universal Access in the Information Society (UAIS), special Issue on "Guidelines, standards, methods and processes for software accessibility", Springer Verlag, Vol.3, N.1, pp. 57-70.
4. Leporini, B., Paternò, F., Scorcia, A., (2006). Flexible Tool Support for Accessibility Evaluation. Interacting with Computers, Elsevier, to appear, 2006.
5. Mariage, C., Vanderdonckt, J., Pribeanu, C. State of the Art of Web Usability Guidelines, Chapter 41, in R.W. Proctor, K.-Ph.L. Vu (Eds), "The Handbook of Human Factors in Web Design", Lawrence Erlbaum Associates, Mahwah, 2005.
6. Mayhew D. J.. Principles and guidelines in software and user interface design. Prentice Hall, Englewood Cliffs, New Jersey, 1992.
7. Nicolle, C., Abascal J. Inclusive design guidelines for HCI, p. 285, Taylor & Francis, 2001.
8. Smith S. L. and Mosier J. N. Guidelines for designing user interface software. Mitre Corporation Report MTR-9420, Mitre Corporation, 1986.
9. Web Accessibility Guidelines 1.0. Web Accessibility Initiative, W3C Recommendation 5-May-1999. Accessible at http://www.w3.org/WAI/GL/WCAG10/
10. Web Content Accessibility Guidelines 2.0, W3C Working Draft 17 January 2006, available at http://www.w3.org/WAI/GL/WCAG20/
11. XML Schema Part 0: Primer Second Edition. W3C Recommendation 28 October 2004 available at http://www.w3.org/TR/xmlschema-0/

Interpreting Results from Large Scale Automatic Evaluation of Web Accessibility

Christian Bühler[1], Helmut Heck[1], Olaf Perlick[1], Annika Nietzio[1], and Nils Ulltveit-Moe[2]

[1] Forschungsinstitut Technologie-Behindertenhilfe (FTB)
der Evangelischen Stiftung Volmarstein, Grundschötteler Str. 40
58300 Wetter (Ruhr), Germany
eiao@ftb-net.de
www.ftb-net.de

[2] Høgskolen i Agder, Fakultet for Teknologi
Grooseveien 36, 4876 Grimstad, Norway
Nils.Ulltveit-Moe@hia.no
http://www.hia.no

Abstract. The large amount of data produced by automatic web accessibility evaluation has to be preprocessed in order to enable disabled users or policy makers to draw meaningful conclusions from the assessment. We study different methods for interpretation and aggregation of the results provided by automatic assessment tools. Current approaches do not meet all the requirements suggested in the literature. Based on the UCAB approach decribed in UWEM 0.5 we develop a new aggregation function targeted at the requirements.

1 Introduction

In the Information Society, where a lot of information is made available on the web, it is essential to make this content accessible to all people including people with disabilities.[1] To get an overview of the accessibility status of a large number of sites manual evaluation by experts or disabled users can produce the most reliable results but often turns out to be too time-consuming and expensive. An automatic assessment of web accessibility is an alternative even though it can not perform all the necessary tests for a conformance claim. However, it can measure certain features that can be utilised as indicators for accessibility and allows the monitoring of a large number of web sites.

The majority of existing automatic web accessibility assessment tools (e. g. Watchfire Bobby,[2] or ATRC web accessibility checker[3]) are *critique systems* designed for the needs of web site developers. The user can select a set of guidelines

[1] The European Union has defined eInclusion as part of the Lisbon strategy: "Ensure that every citizen should have the appropriate skills needed to live and work in a new Information Society for all."

[2] http://webxact.watchfire.com/

[3] http://tile-cridpath.atrc.utoronto.ca/acheck/servlet/ShowGuide

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 184–191, 2006.

like WCAG 1.0 [1]. The tool checks the web site or single web pages against these guidelines and reports a list of error messages often combined with repair suggestions. Such a report is not very helpful for disabled people or policy makers who have questions like

- How accessible is this web site for a certain disability group?
- How accessible is this web site compared to previous versions or compared to other sites?

A conformance approach that summarises the evaluation result into a conformance category (WCAG A, AA or AAA) is too coarse to answer these questions. Therefore, it is useful to present the results instead as a continuous quality measure that allows comparison and grading.

The EIAO project[4] is currently establishing the technical basis for a European Internet Accessibility Observatory (EIAO). An internet robot for automatic and frequent collection of data on web accessibility has been developed. The evaluation of this data is performed by a set of web accessibility metrics that reports accessibility problems and deviations from web standards according to UWEM 0.5 [2] and WCAG 1.0. A data warehouse will provide on-line access to collected accessibility data. To present the large amount of data to the public we need meaningful interpretation and aggregation methods.

The remainder of this paper is organised as follows. In section 2 we discuss the requirements that a large scale assessment of web accessibility should meet. We also review some approaches addressing the interpretation of results from automatic web accessibility evaluation that have been proposed in the literature. The next section presents new aggregation functions for large scale web accessibility evaluation and shows how they comply with the requirements. We conclude with the results from preliminary experimental evaluation. We also point out some open questions and give prospects for future research directions.

2 Measuring Web Accessibility

Determining the accessibility of a web page can be viewed as a two stage process [2]. First, the resources are inspected with regard to possible barrier types. Several different properties (e. g. nesting, relation and attribute values of HTML elements) are extracted and reported (e. g. via an EARL report [3]).

The goal of the second stage is to compute a single value representing the accessibility of the web page from these fine-grained reports. In this paper we examine methods for the second stage.

These values can facilitate the comparison and presentation of the results. Further statistical analyses allow estimation of the accessibility of a whole web site or groupings of web sites from the results for the parts.

[4] The EIAO project is co-funded by the European Commission, under the IST contract 2003-004526-STREP.

2.1 Requirements

In his study of requirements for a web accessiblity metric, Zeng [4] discusses the following properties:

1. Continuous range of values (more discriminative power than binary pass-fail-results or conformance levels)
2. Take into account size and complexity of web site (or web page)
3. Efficient computation (scalability)[5]
4. Normative definition of accessibility (derived from WCAG or other standard)[6]

From UWEM 0.5 [2] two further requirements can be derived:

5. Enable unique interpretability, repeatability and comparability of results
6. Take into account different disability groups

During the developement of methods for the EIAO project it has become apparent that there is an additional requirement for large scale automatic web accessibility assessment:

7. Support for efficient sampling algorithms[7] (provide preliminary results for parts of the web site already during data collection)

However, this is beyond the scope of the paper. We will focus our study on requirements 1, 2, 5, and 6.

2.2 Terms and Notation

We will use the following notation to refer to the quantities of a web resource that are involved in the calculations.

b is the *barrier type* (e. g. the UWEM test name)
u is a *disability group* (e. g. blind, hard of hearing, physically disabled)
i is a unique *identifier of the location* that was inspected (e. g. URL + line/column number or URL + XPath)
A *sample* p is denoted by a set of location identifiers $p = \{i_0, i_1, \ldots, i_n\}$ containing all locations from the relevant web pages, or key use scenarios.
The results from stage one are given by a *report*

$$R_{ib} = \begin{cases} 1 & \text{barrier } b \text{ detected at location } i \\ 0 & \text{no evidence for } b \text{ at location } i \end{cases}$$

[5] This requirement is met because we consider only functions that can be calculated from a closed expression. Methods involving higher level statistical analysis of the results from stage one (e. g. statistical / machine learning approaches) will not be taken into account here.

[6] We assume that the results from stage one have been derived in accordance with WCAG or an other standardised methodology.

[7] Sampling algorithms can be employed to sample to a given error margin within a given confidence interval. This improves performance because a complete assessment of all pages from a web site is not necessary.

The total number of reports for barrier b within sample p is denoted by N_{pb}. $B_{pb} = \#\{i \in p : R_{ib} = 1\}$ denotes the number of fail reports for barrier b. A sum over all barrier types b yields $N_p = \sum_b N_{pb}$ the total number of reports for p and $B_p = \sum_b B_{pb}$ the total number of fail reports for p.

The *severity* of barrier type b for disability group u is given by $S_{ub} \in [0; 1]$.[8]

2.3 Related Work

Sullivan & Matson, 2000. Sullivan and Matson [5] describe the evaluation of eight priority 1 checkpoints from WCAG. They calculate the ratio between the potential points of failure and the actual points of failure:

$$\text{failure rate}(p) = \frac{B_p}{N_p} \tag{1}$$

This approach does not distinguish the barrier types, it only counts the total number of barriers reported B_p.[9] The result is interpreted as failure rate. (0 meaning no accessibility problems and 1 complete failure.)

This approach meets requirements 1 and 5. Requirement 2 is addressed as well. The barrier model does not distinguish different barrier types and consequently does not offer support user group modelling (req. 6).

Zeng, 2004. Zeng [4] proposes a scoring function called "WAB score" for a web page p.

$$\text{WAB}score(p) = \sum_b \frac{B_{pb}}{N_{pb}} w_b \tag{2}$$

where w_b is the inverse of the WCAG priority of the checkpoint relevant for barrier b. A high $WABscore$ means low accessibility. The score for a web site $S = \{p_0, p_1, \ldots, p_m\}$ is the arithmetic mean of the scores for the individual pages.

$$\text{WAB}score(S) = \frac{\sum_{p \in S} \sum_b \frac{B_{pb}}{N_{pb}} w_b}{|S|} \tag{3}$$

This approach complies with requirements 1 and 5. It has no support for different disability groups. The handling of complexity (ratio of encountered violations and possible violations) favours samples with few barrier types. E. g. a page with three true barriers out of three potential barriers will get a score of $3w_b$ if all three barriers have a different type, but only w_b if the three barriers have the same type.

[8] The severity is sometimes denoted by the term *barrier probability* (i. e. the probability that a barrier of type b is a barrier for disability group u.) In UWEM 0.5 the notation F_{cui} is used, where c is a WCAG checkpoint, u a disability group, and i a failure mode (equivalent to barrier type b in our notation).

[9] It is mentioned that the formula has been adapted to include the size of the web page – 10 failures out of 100 should be treated differently from 1 failure out of 10. However, a formal description of this procedure is not given.

UWEM 0.5. The Unified Web Evaluation Methodology (UWEM) 0.5 [2] introduces a probabilistic model: the User Centric Accessibility Barrier model (UCAB). The barriers are grouped by WCAG checkpoint c. The severity is denoted by F_{cub}. Because the barriers are assumed to be (statistically) independent the barrier probability of sample p for disability group u can be calculated by multiplication of the involved probabilities F_{cub} (probability that barrier type b constitutes a barrier for disability group u).

$$F(p,u) = 1 - \prod_{c=1}^{n}(1 - F_{cub}) \tag{4}$$

where n is the number of checkpoints. A lower value of $F(p,u)$ indicates higher accessibility.

This approach satisfies requirement 6 because it allows calculation of assessments for different user groups.[10] It also has properties 1 and 5, but the complexity of the inspected web resource it not taken into account.

The aggregation includes only one barrier b for each checkpoint c. It is not clear how b is determined if there are multiple barriers corresponding to one checkpoint c. As we are looking for a statement about the accessibility of the entire sample p it seems necessary to take into account all barriers that were reported.

Note, that this methodology is still under development. The upcoming version UWEM 1.0 will presumably address some of the issues mentioned above.

Other Methods. Some methodologies, like BITV short test[11]and AIR judging form[12] also perform aggregation of test results. However, they are not subject to our investigation because they are based on manual assessment. The reports contain only one statement about the occurence of each barrier within the whole resource and are not broken down into statements for single locations, which is a prerequisite for the approaches we compare.

3 An Improved Aggregation Method

We base our further exploration of aggregation functions on the UCAB model [2] because of the advantages it has over the simple additive calculation of score values (cf. discussion of Zeng's approach).

First of all we propose to aggregate *all* reports for one sample p (Notation as described in section 2.2).

$$A_1(p,u) = 1 - \prod_{\{R_{ib}:i \in p\}} (1 - R_{ib}S_{ub}) \tag{5}$$

[10] UWEM 0.5 does not cover the question how the severity values for the different user groups can be estimated. We will address this issue in section 3.2.

[11] `http://www.wob11.de/bitvkurztest.html`

[12] `http://www.knowbility.org/`

Subsequently, we address the two major issues: How can complexity of the web resource and needs of different disability groups be taken into account?

3.1 Including Complexity

The function A_1 can be converted into a product over barrier types

$$A_2(p,u) = 1 - \prod_b (1 - S_{ub})^{B_{pb}} \tag{6}$$

This is easy to verify as the number of factors $(1 - S_{ub})$ for one barrier b is B_{pb}.

Note that this formula enables modelling of *absolute* barriers. A barrier is absolute if it prevents to user from completing a task. The severity for such a barrier is $S_{ub} = 1$. This yields $A_2(p,u) = 1$ because the product becomes zero (one of the factors is $(1 - S_{ub})^{B_{pb}} = (1-1))^{B_{pb}} = 0)$.

We propose to model the complexity of the sample by adapting the exponent in this formula.

$$A_3(p,u) = 1 - \prod_b (1 - S_{ub})^{C_{pb}} \tag{7}$$

where C_{pb} is a value describing the complexity of p with regard to barrier type b. Quantities relevant for the calculation of C_{pb} are B_{pb}, N_{pb} and, N_p.
C_{pb} should satisfy

- If no barriers of type b are encountered there is no contribution to the aggregation function. $(B_{pb} = 0 \Rightarrow C_{pb} = 0)$
- If a barrier of type b is encountered this will decrease the result of the aggregation function. $(B_{pb} > 0 \Rightarrow C_{pb} > 0)$

The following complexity function has the desired properties.

$$C_{pb} = \frac{B_{pb}}{N_{pb}} + \frac{B_{pb}}{B_p} \tag{8}$$

This function takes into account the ratio of potential and actual barriers. And in addition the ratio of all failures to the number of failures for one barrier type. This additional contribution ensures that barriers are considered according to their overall proportion of occurrences within the web resource.

3.2 Estimating Severity

Simple Heuristics. The barrier types which are relevant for a specific disability group can be identified rather straightforward. E. g. for deaf users without visual impairment a missing textual description of an image is not a barrier.

This yields the following estimate (all relevant barrier types get the same weight).

$$S_{ub} = \begin{cases} 0 & \text{barrier type } b \text{ is not relevant for disability group } u \\ s > 0 & \text{barrier type } b \text{ is a barrier for disability group } u \end{cases}$$

User Model. In an iterative process the results from automatic evaluation and aggregation are compared to results acquired from manual testing. The initial values are adapted accordingly to improve the predictions of the automatic system.

Involving Disabled Users. The most reliable way of estimating severity weights involves feedback from disabled users. I. e. asking users to rate the severity of the different barrier types.

4 Experimental Evaluation

The goal of an aggregation function is to provide a value that indicates the accessibility of a web page for a disabled user. Therefore we chose to compare the results from automatic web accessibility assessment with the ratings given by a group of fifteen disabled people during a user testing study.

Table 1. Results from experimental evaluation

	page A		page B		page C		page D		page E		page F	
	all	blind	all	blind	all	blind	all	blind	all	blind	all	blind
Sullivan & Matson	0.33	n/a	**0.20**	n/a	0.34	n/a	**0.08**	n/a	**0.57**	n/a	**0.27**	n/a
UWEM 0.5	0.99	0.99	0.51	0.51	0.60	0.60	0.97	0.97	0.10	0.10	0.96	0.96
A_3	**0.42**	**0.42**	**0.20**	**0.20**	**0.15**	**0.15**	0.10	**0.10**	0.15	0.15	0.36	**0.36**
User rating	0.58	0.58	0.30	0.13	0.10	0.08	0.33	0.08	0.47	0.75	0.29	0.38

Table 1 presents the average results form all users and results for a selected user group. The reports that are the input to the aggregation functions were generated by the EIAO observatory.[13] The severity values are set to $S_{ub} = 0.05$ for all barrier types b and disabled user groups u.

The evaluation shows that the values from the improved aggregation function A_3 in most cases are closest to the user ratings. It is interesting to see that the simple failure rate measure proposed by Sullivan & Matson also yields good results.[14]

5 Conclusion and Research Prospects

We studied different methods for interpreting results from large scale automatic evaluation of web asseccibility. We found that current approaches do not meet all the requirements suggested in the literature. Based on the UCAB approach

[13] EIAO covers automatically testable features that are a subset of the WCAG 1.0 AA requirements.

[14] The aggregation method proposed by Zeng has not been included in the comparison because the WAB score has no upper boundary and can therefore not be normalised to values in $[0; 1]$.

decribed in UWEM 0.5 [2] we developed a new aggregation function targeted at the requirements. A preliminary experimental evaluation shows some promising results. To strengthen the statistical evidence additional experiments involving more users will be conducted. There are still many open questions in this field. Directions for future research include:

- *Improved modelling of key use scenarios:* A key use scenario is sequence of tasks that a user performs on a web site. An accessibility evaluation has to take into account that there are crucial parts within the scenario (following links, filling in forms) which should be modelled accordingly.
- *Introducing uncertain reports:* Some barriers can only be identified with limited confidence. The range of the reports can be extended to probability values $R_{ib} \in [0; 1]$.

Acknowledgements

The authors would like to thank Jenny Craven and Peter Brophy who designed and conducted the user testing experiments and all other EIAO partners who provided comments and valuable feedback during our work.

References

1. W3 Consortium: Web content accessibility guidelines 1.0. Available at `http://www.w3.org/TR/WCAG10/` (1999)
2. Web Accessibility Benchmarking Cluster: D-WAB2 unified web evaluation methodology (uwem 0.5). available from `http://www.wabcluster.org/uwem05/` (2005)
3. McCathieNevile, C., Abou-Zahra, S.: Evaluation and report language (EARL) 1.0 Schema. W3C editor's working draft. Available at `http://www.w3.org/WAI/ER/EARL10/WD-EARL10-Schema-20060101` (2006)
4. Zeng, X.: Evaluation and Enhancement of Web Content Accessibility for Persons with Disabilities. PhD thesis, University of Pittsburgh (2004)
5. Sullivan, T., Matson, R.: Barriers to use: Usability and content accessibility on the web's most popular sites. In: Proceedings of ACM Conference on Universal Usability – CUU. (2000)

Managing and Monitoring Web Site Accessibility

Shadi Abou-Zahra

W3C and Web Accessibility Initiative
2004, Route des Lucioles BP 93 – 06902 Sophia-Antipolis, France
shadi@w3.org
http://www.w3.org/People/shadi/

Abstract. Evaluating Web sites for accessibility is a quality assurance process that becomes increasingly difficult to manage as the size and complexity of a Web site increases. There is a growing need to effectively manage and monitor the accessibility of Web sites throughout the development process. The Evaluation and Report Language (EARL) is a semantic Web vocabulary, and is a W3C royalty-free format for expressing test results. While EARL can be used to support generic Web quality assurance processes, it has been specifically developed to assist Web accessibility evaluation reviews. EARL facilitates the exchange of test results between development and quality assurance tools. While a lot of development has already gone into EARL and tools that support the EARL format, a lot more work and research remains to be addressed.

1 Introduction

While the markup code of Web sites can generally be automatically validated against formal grammars, the accessibility features are mostly evaluated with manual input from human reviewers. For example, it is generally not possibly to automatically verify the validity of a textual description for an image[1]. Depending on the size and complexity of a Web site as well as on the type and thoroughness of the evaluation review, a considerable amount of time and resources may be required to determine the level of accessibility. Also, significant amounts of evaluation test results can be generated from detailed evaluations of Web sites for accessibility.

The challenge for many Web site owners is to employ quality assurance processes for Web accessibility that will improve the efficiency of evaluation reviews as well as enhance the management of issues. This paper describes how the Evaluation and Report Language (EARL) [1], a common format for expressing test results, supports such quality assurance processes. This paper will highlight some of the key features of EARL that facilitate monitoring and managing the accessibility of Web sites.

1 With the exception of images of text that can sometimes be automatically processed using character recognition. However, this approach has many limitations and often does not work.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 192 – 198, 2006.

2 Types of Web Accessibility Evaluation Reviews

There are many of types of Web accessibility evaluation reviews that range from less formal preliminary reviews to more comprehensive approaches that include technical and user testing of accessibility features. The W3C/WAI resource suite called "Evaluating Web Sites for Accessibility" [2] describes some of these approaches in more detail. This paper focuses on reviews that evaluate the conformance of Web sites to accessibility standards such as the Web Content Accessibility Guidelines (WCAG) 1.0 [3]. Other examples of Web accessibility standards[2] include the requirements defined by Section 508 in the USA, BITV in Germany, or JIS in Japan.

While the difference between some of the currently existing standards for Web accessibility may vary significantly, most share a common anatomy[3]. Basically, it can be assumed that there is a target set of criteria to which a Web site is evaluated. The conformance to each criteria is determined by conducting a series of atomic tests. While some of these tests can be executed automatically by software, most of them usually need human reviewers to determine the result of such tests. The sequence in which the tests are executed usually depends on the results from previously executed tests. This introduces a complexity vs. transparency dilemma for reporting evaluation results: if each executed atomic test is recorded then huge reports will be generated but if they are excluded, it becomes less transparent why a criteria was not met.

Another aspect that affects the evaluation review is the sampling strategy. Usually it is not economically feasible to manually evaluate every page on a Web site. Web applications that generate content dynamically also add another dimension of size complexity. Therefore, most Web accessibility evaluation reviews employ some sort of sampling mechanism to reduce the number of tests that need to be manually executed. Basic sampling mechanisms rely on a selection of Web pages that cover different types of features, templates, production methods, or content styles. More sophisticated mechanisms consider additional parameters such as link paths and transactions, page traffic and relative importance, as well as other factors.

3 Role of Evaluation Tools During Review Processes

Despite the fact that most of the tests required for conformance evaluations need to be executed manually, tools can assist many tasks and hence significantly improve the efficiency of Web accessibility evaluation processes. For example, tools can guide reviewers through the testing procedure, highlight issues that may be more applicable to specific areas of the Web site, or provide functionality to help reviewers determine the result of specific tests. W3C/WAI maintains a list of Web accessibility evaluation tools [4] which can be used during evaluation and development processes.

---

2 W3C/WAI maintains a list of international policies relating to Web accessibility [8].

3 Possibly because many of the existing Web accessibility standards are derived (directly or indirectly) from the W3C Web Content Accessibility Guidelines 1.0 [3].

3.1 Tool Support for Automatic Testing

For accessibility tests that can be executed automatically, automation is generally cost effective. Testing can often be repeated periodically on large numbers of pages. Even though the accuracy of Web accessibility evaluation tools may vary, the error rate per tool and per Web site can be usually assumed as constant and thus simple to adjust. Also, it is important to note that while some tests can not be executed automatically, sometimes their applicability on a given Web page or site can be automatically determined. This potentially reduces the amount of manual evaluations that need to be carried out by human reviewers and improves the efficiency of evaluations.

3.2 Tool Support for Manual Testing

Web accessibility evaluation tools can support human reviewers to manually carry out accessibility tests in a number of ways. So called *transformation tools* modify the presentation of Web pages to help reviewers find potential barriers for people with disabilities. For example, transformation tools could display Web pages in low color contrast, with large font sizes, or simulate page elements as they would be presented by assistive technology[4]. Other functions of evaluation tools can highlight areas of the Web pages or the underlying markup code (such as HTML or CSS, etc.) to help reviewers identify barriers of more technical nature on a Web site.

4 Need for a Common Format for Evaluation Results

Some enterprise Web accessibility evaluation tools and custom solutions implement different formats for expressing evaluation test results. This enables these types of tools to integrate different modules (or other evaluation tools) for automatic and manual evaluations into a more complete testing framework[5]. Such frameworks often offer other data monitoring and analysis capabilities. For example, visualizing the number of issues encountered on a Web site against time to monitor the performance over time. Also, some authoring tools (such as editors or content management systems) provide APIs[6] that can be used to integrate the output from evaluation tools.

This exchange of test results between different types of development tools generally contributes to more efficient development of accessible Web sites. However, only few tools provide support for the exchange of test results. The reason for this rather slow implementation of support is possibly the cost of developing a format for expressing test results, or for developing the API functionality to process such data. Also the large amount of different and mostly proprietary formats and APIs creates a challenge for developers of evaluation tools to support them effectively.

One possible approach to address this lack of support for exchanging results between different types of development tools is to provide a commonly accepted format for expressing evaluation test results. Such a standardized format would encourage evaluation tool developers to provide mechanisms for exporting test results in this

4 Assistive technology are computer programs or devices that assist users with disabilities accessing the Web. These include screen readers, screen magnifiers, or voice command tools.
5 Examples include HiSoftware AccVerify, Watchfire WebXM, or WebThing AccessValet.
6 Examples include Macromedia Dreamweaver, Microsoft Front Page, or RedDot XCMS.

format because it can be potentially processed by a large number of tools. Conversely, authoring tool developers would be encouraged to provide mechanisms for importing evaluation results in this standardized format as it allows the integration of potentially more variety of evaluation tools. Finally, third-party data analysis tools (for example to prioritize the repair of reported problems according to their cost of repair or impact on Web accessibility; or to generate customized reports for specific developers) can be developed based upon this commonly used format for expressing test results.

5 Features of the Evaluation and Report Language

The Evaluation and Report Language (EARL) [1] is a simple vocabulary to record the following aspects of a quality assurance test:

- Which content was tested?
- Which criteria was it tested for?
- Who/what carried out the test?
- What is the outcome of the test?

While these questions apply to generic Web quality assurance testing (for example to determine conformance against usability or corporate design guidelines), many of the use cases and scenarios for developing EARL have originated from requirements of Web accessibility evaluation reviews. EARL can also be used beyond quality assurance processes on the Web, for example for generic software testing, but this remains outside the scope of this paper.

The vocabulary of EARL has been developed using the W3C "Resource Description Framework" (RDF) [5] which is a semantic Web format that allows the definition of custom vocabulary in a machine readable form. EARL is being developed under the W3C process[7] which encourages the participation of different stakeholder and ensures consensus among them. W3C specifications are also royalty-free.

Another compelling reason to implement the EARL vocabulary using RDF is to benefit from the existing and growing semantic Web community. For example, EARL reports can be queried using SPARQL[8]. RDF also finds support in databases and APIs in different programming languages[9]. Furthermore, RDF is flexible and enables languages such as EARL to be extended or refined for specific usages yet retain the overall structure and compatibility between tools. For example, developers could adopt different test criteria descriptions yet retain interoperability of the reports.

Use Cases. The following are brief descriptions of some of the use cases for a common format for expressing test results in the context of evaluating Web sites for accessibility.

Combining Results. A standardized format for expressing test results allows the combination reports from different Web accessibility evaluation tools. This enables different reviewers to carry out separate evaluation tasks using different tools, or for

[7] The W3C Process Document is available at `http://www.w3.org/Consortium/Process/`

[8] SPARQL is an SQL-like query language for RDF. Other examples include RDQL or Squish.

[9] For example Jena for Java, RAP for PHP, or Raptor for Python (and other languages).

reviewers to employ different tools at different stages evaluation process and combine the results.

Comparing Test Results. Calibrated results, such as results from pre-defined test suites could be recorded in a standardized format to facilitate the benchmarking of Web accessibility evaluation tools. This allows tool users to compare the output from different tools, and for tool developers to better test and improve their tool implementation.

Processing Results. Quality assurance tools can rely on a standardized format to analyze, sort, prioritize, or infer conformance claims. Such processing tools could also provide their output in the same initial format to allow cascades of processing tools. For example, one tool could specialize in collecting results, while another one prioritizes them.

Generating Reports. Tools that specialize in reporting could provide customized views on the results so as to provide specific developers with information according to their preferences. For example, Web programmers may want to receive more verbose bug reports with line numbers and error messages, while project managers and executive may want to receive higher level management reports.

Integration into Authoring Tools. As already described, a standardized format for expressing test results provides a mechanism for the integration of Web accessibility evaluation tools into authoring tools. It also allows evaluation tool developers to focus on testing techniques and algorithms and relying on authoring tools to provide the interface to the developers.

Integration into Web Browsers. Web accessibility evaluation results could be used by Web browsers to improve the experience of the end-users. For example, Web browsers could use the results from Web accessibility evaluation tools to detect and linearize complex tables or to suppress moving content according to user preferences.

Integration into Search Engines. Similar to Web browsers, search engines could enhance the experience of their users by considering user preferences. For example, a search engine may be configured to only return Web pages that match the request and that provide certain accessibility features. For example, Web pages that claim to be operable by keyboard[10].

Justifying Conformance Claims. To provide more credibility and transparency, Web site owners may chose to publish some of the test results that led to the conformance claims. For example, to supplement a quality mark label with a consistent report of what has been tested and when, could help the user set expectations about the accessibility level of a Web site.

---

[10] Note that this raises issues related to trust and content labeling. For example, one thought is that third party services may provide information about Web sites rather than Web site owners. However, this discussion is beyond the scope of this paper.

6 Current Status and Future Development

At the time of writing this paper, the Evaluation and Report Language (EARL) [1] is a W3C Working Draft. However, it is fairly mature and expected to proceed to Last Call stage in due course. At the same time, building blocks on which EARL is dependent (such as *Dublin Core* or *Friend Of A Friend* vocabularies, *OWL Web Ontology Language*, or various query languages as described in the previous section) are widely deployed and implemented. So, while EARL itself is still in draft stage, it is gaining stability through the evolution of the related technologies.

This is also reflected by the availability of tools that support the EARL format. There is a wide range of EARL implementations, however most of these produce EARL output and only few tools process EARL reports. It is interesting to see the uptake of EARL in both reference prototypes and research projects[11], as well as in operational tools[12]. The W3C/WAI Evaluation and Repair Tools Working Group (ERT WG) [6] develops EARL, and maintains a list of Resources related to EARL [7]. This list references more implementations and projects.

The main objective for the ERT WG [6] is to resolve the currently outstanding issues in the EARL working draft and publish the first version as a W3C Recommendation. One of the issues includes describing the occurrence of test results (for example accessibility violations) within a Web site. To promote the deployment of EARL, the working group is also developing an EARL 1.0 Guide to complement the EARL 1.0 Schema with more examples and guidance for implementing EARL. The EARL 1.0 specification is therefore becoming increasingly comprehensive yet modularized to fit the needs of different audiences.

At the same time, feedback from implementation experience will be essential to avoid unexpected issues. For example, current implementations show that methods need to be developed that reduce the amount of output generated by EARL reports. Also, there is a strong need to increase the uptake of EARL in authoring tools, especially in content management systems. While many of these implementations could be reference prototypes and research projects during the development of EARL, deployment in operational tools will be necessary when EARL become more mature.

There are also some research questions open, mainly for later versions of EARL. For example, how can EARL reports be more resilient towards changes made on an already evaluated Web site? One approach to address this issue could be by analyzing where the changes occurred on a page and infer which tests a given change may have affected. Another open research question is how to map or relate tests that have been developed by different vendors and that evaluate the same criteria to each other. This is related to the issue that many vendors do not want to disclose the proprietary tests, or the sequence in which they were carried out in order to determine a result.

7 Conclusion

Managing and monitoring the accessibility level of Web sites is a challenging quality assurance task. Several evaluation tool developers have already reacted to this

[11] Examples include EARL filter, EARL client and database, or WAINU.

[12] Examples include HiSoftware AccVerify, Sidar Hera, or WebThing Site Valet.

demand and developed systems to support reviewers or whole review teams in carrying out specialized actions in a comprehensive quality assurance process.

However, a common approach using a standardized format for expressing test results would facilitate the integration of different type of tools. A common standard will promote the implementation into currently available tools as well as the development of new ones. It will simplify data analysis and will enable many different use cases.

The Evaluation and Report Language (EARL) is currently being developed by the W3C/WAI Evaluation and Repair Tools Working Group (ERT WG). It addresses several use cases for integrating Web accessibility evaluation reviews into the development processes. EARL is developed under the W3C Process and is the result of consensus among different stakeholders.

While EARL is basically a simple vocabulary to describe evaluation test results, it has the power and potential for different tasks. EARL uses the Semantic Web Resource Description Framework (RDF) to define its vocabulary. This allows it to be machine readable and to benefit from a rapidly growing community. It benefits from readily available support in databases, query languages, and programming language APIs.

At the time of writing this paper, EARL 1.0 is still a W3C Working Draft but relatively mature. There is some existing tools support but also the opportunity for much more research and development work.

References

1. Evaluation and Report Language (EARL) `http://www.w3.org/WAI/intro/earl`
2. Evaluating Web Sites for Accessibility `http://www.w3.org/WAI/eval/`
3. Web Content Accessibility Guidelines (WCAG) `http://www.w3.org/WAI/intro/wcag`
4. List of Web Accessibility Evaluation Tools `http://www.w3.org/WAI/ER/tools/`
5. Resource Description Framework (RDF) `http://www.w3.org/RDF/`
6. Evaluation and Repair Tools Working Group (ERT WG) `http://www.w3.org/WAI/ER/`
7. EARL Resources `http://www.w3.org/WAI/ER/earl`
8. Policies Related to Web Accessibility `http://www.w3.org/WAI/Policy/`

Semi-automatic Evaluation of Web Accessibility with HERA 2.0

Carlos Benavídez[1], José L. Fuertes[2], Emmanuelle Gutiérrez[1], and Loïc Martínez[2]

[1] Sidar Foundation. Spain
{carlos, emmanuelle}@sidar.org
[2] Universidad Politécnica de Madrid. Spain
{jfuertes, loic}@fi.upm.es

Abstract. The evaluation of the accessibility of a web site calls for the participation of human evaluators: most of the checkpoints to be assessed cannot be evaluated fully automatically. This paper presents the second version of HERA, a multilingual online tool developed by the Sidar Foundation that automatically performs a preliminary analysis of a web page and then provides support for the complete manual evaluation process. This description includes the justification for a newer version, the technologies used, and the main strengths of HERA 2.0 as compared with other tools and version 1.0.

1 Introduction

In the last few years the issue of web accessibility has gained in relevance all over the world [1,2]. In the European Union, for instance, most member states have developed legislation to ensure the accessibility of public administration web sites at all levels (central, regional, local…) in conformance with the World Wide Web Consortium's Web Content Accessibility Guidelines 1.0 (WCAG) [3].

In this context, the evaluation of the accessibility of a web site is of utmost importance. This evaluation cannot be completely automated, as many of the checkpoints require human judgement to assess a web page's conformance level. Thus, the evaluation of web accessibility is a complex task requiring human expertise and tool support [4,5]. The person performing this task needs sound knowledge and experience in web development and has to be proficient in the use of the techniques required to evaluate conformance for each of the WCAG 1.0 checkpoints.

Therefore, both experts and novices in web accessibility evaluation have a common need: a tool that provides support for the manual evaluation of web accessibility, automating as much of the work as is possible.

This tool should provide services for all kinds of users. Firstly, the tool should provide rapid access to the text of the checkpoints and guidelines, because even experts find it difficult to correctly remember all 14 guidelines and 65 checkpoints. Secondly, the tool should provide information about all the items that require evaluation for each checkpoint, because some of the checkpoints refer to several structural elements. In third place, the tool should provide means to improve the visual detection of problems in the web page or in the source code by automatically

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 199 – 206, 2006.

highlighting the different items on the page. In addition, the tool should provide services for the users to store either complete or incomplete evaluation results for further reference. Finally, the tool should automate as much work as possible, such as detecting the checkpoints that are not applicable to any one web page, the checkpoints that are correctly implemented and others that are wrong.

With these goals in mind, the Sidar Foundation [6], in cooperation with the Technical University of Madrid (UPM), has developed the HERA system, an online tool for the semi-automated evaluation of the level of conformance with the checkpoints of WCAG 1.0.

This paper presents ongoing work on the second version of HERA [7], [8]. Section 2 will briefly discuss related work. Section 3 will present a justification of the need for a second version of HERA. Section 4 will give a description of the main technical issues of the current version: automatic preliminary evaluation, use of PHP instead of CSS and multiple format report generation. Finally, section 5 will present some concluding remarks about the experiences in the use of HERA, along with work to be undertaken in the future.

2 Related Work

There are many automatic and semi-automatic evaluation tools that can be used to analyse a web page and generate an assessment of its accessibility level [9]. In this paper we will focus only on free online tools, as this is the category into which HERA falls. Three of the most relevant tools are listed below:

- Cynthia Says [10], from the International Centre for Disability Resources on the Internet (ICDRI), in collaboration with Hi-Software. This is an automated accessibility checker that can be used to test one web page per minute. It can generate a report based on section 508 standards (US legislation [11]) or on WCAG checkpoints, with an additional evaluation of the quality of alternative texts. This is an exclusively automatic tool that does not check all of the checkpoints and provides no support, beyond a checklist, for manual evaluation. It is only available in English.
- TAW [12], from the Centre for the Development of Information and Communication Technologies (Spain), is an online service that evaluates web page accessibility. It provides a report showing whether or not some checkpoints are satisfied, and displays tags placed in items that should be checked manually. This tool is available in English and Spanish. Again, the online tool does not support manual evaluation, although there is a stand-alone freely downloadable version that does.
- WebXACT [13], from Watchfire, is a free online service. It tests only some of the WCAG checkpoints, but suggests that the user should test some others. In addition, WebXACT also performs other tests, such as quality measures, speed information, etc. This tool was previously known as Bobby (name now used for a commercial stand-alone product), and it is only available in English.

Fully automatic tools like the ones above clearly inform the user about the limitations of an automated evaluation and emphasize possible problems that have to be manually verified by the human inspector. Moreover, they usually provide additional information and examples on how to use items properly to make them more accessible. However, non-expert users actually find it difficult to evaluate these items manually because they do not usually have precise and clear instructions about how to deal with each case in the web page under evaluation.

When comparing HERA with the above tools, there are some relevant differences, as shown in Table 1. They all perform an automatic analysis of as many checkpoints as possible, but HERA is the only online tool that supports manual verification (providing extensive help, modified views of the web page for the evaluation of some checkpoints and storage of evaluation scores and commentaries) and report generation (both human-readable and in the EARL language [14]).

Table 1. Comparison of free web accessibility evaluation tools

Tool	Online	Automatic Analysis	Help on techniques	Modified views	Results storage	Reports
Cynthia Says	Yes	Yes	No	No	No	Human
TAW online	Yes	Yes	No	No	No	Human
TAW stand-alone	No	Yes	Yes	Yes	Yes	Human/EARL
WebXACT	Yes	Yes	No	No	No	Human
HERA	Yes	Yes	Yes	Yes	Yes	Human/EARL

3 Need for HERA 2.0

HERA version 1.0 [15] was made freely available to the public in 2003 and was a great success, as has been demonstrated by the positive feedback that we have received from system users. It is an online tool, based on PHP [16] and CSS [17], that provides support for the manual evaluation of the accessibility of web pages. It was the first online tool for accessibility evaluation that was developed with the manual evaluation process in mind. This tool provides help about the meaning of each of the checkpoints, instructions for the manual evaluation of the checkpoints, a modified view of the page using CSS, a view of the source code of the page, a system to annotate the results of the evaluation of each checkpoint and, finally, a report generation tool (Figure 1).

This first version of HERA used a set of style sheets written in CSS in order to identify and highlight specific web page items, allowing the evaluator to examine the different items and their properties directly by means of a specialised view of a web page, without having to inspect the source code. In doing so, there is no need to modify the original page, except for the minimal change of including a reference to the required style sheet for each checkpoint.

However, we have found that there is a major obstacle to the use of CSS for accessibility evaluation. This is because most browsers provide no support for some of the advanced properties of CSS2 [17], like generated content and attribute value-based rule selectors. In fact, only the Opera browser [18] correctly interprets all the

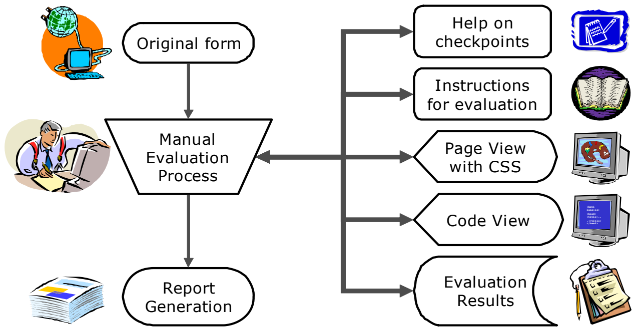


Fig. 1. Overview of HERA version 1.0

style sheets developed for HERA, though we expect more user agents to provide the required level of support for CSS2 in the future.

There are two main reasons that justify the development of a new version of HERA: one is technical and related to questions of implementation, and the other is based on usability (efficiency) issues.

The technical reason for developing HERA 2.0 is the above-mentioned dependence of HERA 1.0 on some advanced elements of CSS2 that today are only fully rendered by the Opera browser. Many users of HERA complained about the fact that they could not use their preferred browser and had to use a different one, with which they were less well acquainted. In some cases this was an important issue, such as for disabled users using specific assistive technologies that did not work well with the Opera browser. So it was decided that HERA 2.0 should be browser-independent to accommodate the maximum range of potential users.

The usability-related reason for HERA 2.0 is efficiency. Performing a complete manual evaluation of the accessibility of a web page takes a considerable amount of time. An inspector using HERA 1.0 had to manually evaluate each of the 65 checkpoints of the WCAG 1.0, even when some of them were not applicable (for example, table-related checkpoints in a web page without tables) or had a value (positive or negative) that would be easy to assign automatically. In HERA 2.0 we decided to reduce the minimum time for evaluating one web page by incorporating a preliminary automated analysis.

4 HERA 2.0

To deal with these new requirements, changes had to be made to the design and implementation of HERA 2.0. As shown in Figure 2, there are three main changes in the new version: an automatic preliminary analysis process (with an accompanying result summary for navigation), the generation of modified page views that does not rely on browser-specific features and, finally, an enhancement of the report generation functionality.

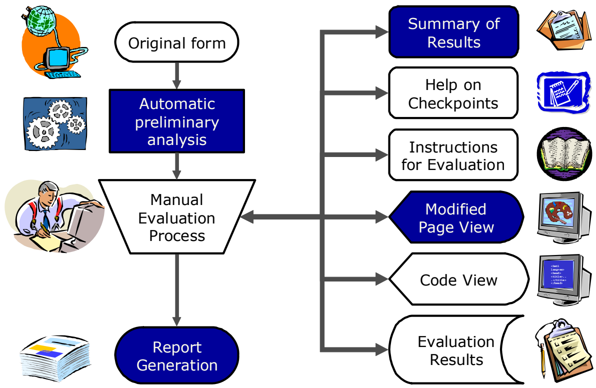


Fig. 2. Overview of HERA 2.0. Main changes from version 1.0 are highlighted.

4.1 Automatic Preliminary Analysis

The first change was the design and implementation of an automatic preliminary analysis, which was programmed using PHP. This process inspects the web page and automatically assigns a value for each of the 65 checkpoints:

- *Pass*: the web page conforms to the checkpoint. This result can only be yielded for a very small subset of the WCAG. For example, when the web page conforms to a formal grammar of the mark-up language used (HTML, XHTML…).
- *Fail*: the web page does not conform to the checkpoint. This result can be output on a larger set of the checkpoints. For instance, when an image element does not have an "alt" attribute.
- *Not applicable*: the checkpoint is not applicable to the analysed web page. For example, if the checkpoint is about frames and the page does not use this technology.
- *Needs checking*: the tool cannot decide and the user is asked to manually evaluate the checkpoint.

It has to be noted that HERA 2.0 seamlessly integrates in its user interface the markup validation services offered by the W3C to check the correctness of HTML, XHTML and CSS.

Table 2 shows a summary of the checkpoints that can be automatically evaluated by HERA. There are some checkpoints (or parts of them) that can be automatically checked, some that should be manually checked and some others on which, although the tool can do some automatic testing, users should make a final decision based on their knowledge and the pages generated by HERA. This means that it is not always possible to perform a fully automated evaluation for some checkpoints and the human inspector should finish off the evaluation of these checkpoints manually.

Several checkpoints appear in more than one column in Table 2. For instance, checkpoint 1.1 (text alternatives) can be automatically evaluated as "fail" if there are images without the "alt" attribute, as "non applicable" if there are no non-textual elements (images, objects, etc.) in the page, and as "needs checking" if there are images with “alt” attributes that require human evaluation to assess whether the alternative text matches the image. On the other hand, this checkpoint will never be automatically evaluated as “pass”, because the computer cannot judge the adequacy of alternative texts.

Table 2. Summary of checkpoints automatically analysed by HERA

	Pass	Fail	N/A	Human
Priority 1	2	4	12	16
Priority 2	17	18	9	21
Priority 3	6	6	6	13
Total	**25**	**28**	**27**	**50**

Once the preliminary analysis has been completed, HERA shows the results of this process, including information such as the URL under analysis, time spent on the automatic evaluation, number of errors found and, more importantly, a table with the summary of the checkpoints evaluation results. This table shows, for each priority level, the number of checkpoints that require manual evaluation, that have passed, that have failed and that are not applicable. This table is used as the main navigation tool during the manual evaluation process.

It should be noted that, although the tool performs an initial evaluation and decides whether some checkpoints pass, fail or are not applicable, the responsibility ultimately lies with the human evaluator. The user should also check the decisions made by the tool to verify that the result output automatically is absolutely correct. The point here is that, due to the diversity of use (or misuse) of mark-up languages by developers or of scripts for generating dynamic content or changing items, a false positive or negative value could be output in some exceptional cases if the intended semantics is not well implemented. But this usually takes no more than a few seconds to check.

4.2 Browser-Independent Modified Page Views

The second change was the technology used to highlight the items of the pages that have to be analysed for each checkpoint. Instead of using CSS, the HERA 2.0 program generates a new version of the analysed page that contains new items that provide the information needed by the evaluator.

To implement this functionality in a browser-independent way, we needed to use server-side technologies. There are many options in this field: PHP, Java Server Pages (JSP), Perl, Active Server Pages (ASP), and so on.

We decided that, to reduce the implementation effort, the best solution would be a scripting language that could be easily integrated with HTML and provided methods to easily manage texts, because HERA has to analyse document structure and insert some additional code in the pages. PHP was selected because it is an open source technology that works well with Apache servers and, last but not least, because we found it is easier to use than other valid technologies, like Perl or ASP.

Therefore, the modified page view module has been implemented by a set of PHP filters that are launched when the user manually evaluates each of the checkpoints.

4.3 Enhanced Report Generation

HERA 2.0 provides an enhanced report generation module, which offers improved functionality:

- The possibility of downloading the XHTML version of the report. In HERA 1.0 the user had to manually download the XHTML report using his or her browser commands.
- One more output format (PDF). This format is useful when the evaluator wants to publish or send a report that is not easy to modify.
- An improved, more up-to-date EARL version of the report, using the latest version of the language.

5 Conclusions and Future Work

A beta version of HERA 2.0 has recently been released and it is available in 4 languages (Spanish, English, Portuguese and Italian) and it is being translated into a few more (including French, Catalan, Galician, Rumanian and Serb). The feedback that we have received so far from both users that already knew HERA and new users has been extremely positive.

Users already familiar with HERA have found that the new version enables them to reduce the time it takes to perform the evaluation of a web page thanks to the new automatic preliminary analysis offered by the tool. Comparing HERA 1.0 and HERA 2.0, we can say that the time saving is around 40 to 50%. Now it is possible to perform a full evaluation of a web page and to generate a detailed report in about 2 to 3 hours (depending on web page complexity, of course). As compared with TAW, the time it takes to evaluate a web page is reduced by about 20 to 30%. Other online tools (Cynthia Says, WebXACT...) cannot be compared with HERA, because they do not provide support for running a full (manual) accessibility evaluation.

In the case of new users, they have learned how a tool can help them to perform a manual evaluation of web accessibility and how cautious one has to be when dealing with the results of an automated evaluation. Moreover, both the Sidar Foundation and the UPM are using the new version of HERA in their web accessibility courses with very encouraging results.

However, there are some relevant issues that require further work to get HERA to offer a better and more comprehensive service, such as the following:

- There is a clear need for a downloadable executable version of HERA, which would enable the evaluation of local web pages and whole web sites.
- Complex web sites are usually evaluated by teams of reviewers, and HERA should provide support for coordination and communication between them.
- Another very useful functionality for evaluator groups would be a system for measuring the reliability of the evaluators.

References

1. Political Declaration of the “II World Summit of Cities and Local Authorities on the Information Society”. Bilbao, Spain. November, 11 (2005) (URL: http://www.it4all-bilbao.org/declaracion)
2. “Tunis Commitment”. World Summit on the Information Society. Document WSIS-05/TUNIS/DOC/7-E. Tunis. November, 18 (2005) (URL: http://www.itu.int/wsis/docs2/tunis/off/7.html)
3. Chisholm, W., Vanderheiden, G., Jacobs, I.: Web Content Accessibility Guidelines 1.0. World Wide Web Consortium Recommendation (1999) (URL: http://www.w3.org/TR/WCAG10/)
4. Education and Outreach Working Group: Evaluating Web Sites for Accessibility: Overview. World Wide Web Consortium (2005) (URL: http://www.w3.org/WAI/eval)
5. Slatin, J. M., Rush, S.: Maximum Accessibility: Making Your Web Site More Usable for Everyone. Addison Wesley Professional, Boston (2003)
6. Sidar Foundation. Homepage (in Spanish) (2006) (URL: http://www.sidar.org/)
7. Sidar Foundation: Hera tool. Version 2.0 beta (2005) (URL: http://www.sidar.org/hera/index.php.en)
8. Benavidez, C., Gutiérrez, E.: HERA una herramienta para la revisión manual de la accesibilidad web (in Spanish). In: Jornadas de Accesibilidad y Nuevas Tecnologías (JANT 2004). Bilbao, Spain (2004)
9. Abou-Zahra, S., Chisholm, W., Brewer, J.: Evaluation, Repair, and Transformation Tools for Web Content Accessibility. World Wide Web Consortium (2005) (URL: http://www.w3.org/WAI/ER/existingtools.html)
10. ICDRI: Test your site now with Cynthia Says to see if it is accessible! International Center for Disability Resources on the Internet (2006) (URL: http://www.icdri.org/test_your_site_now.htm)
11. Access Board: Electronic and Information Technology Accessibility Standards (for section 508 of the Rehabilitation Act). USA Federal Register. Vol. 65, No. 246. Thursday, December 21 (2000) (URL: http://www.access-board.gov/sec508/standards.htm)
12. CTIC Foundation: “TAW: Web Accessibility Test”. Centre for the Development of Information and Communication Technologies. Spain (2006) (URL: http://www.tawdis.net/taw3/cms/en)
13. Watchfire: Watchfire WebXACT. Watchfire (2004) (URL: http://webxact.watchfire.com/)
14. McCathieNevile, C., Abou-Zahra, S.: Evaluation and Report Language (EARL) 1.0 Schema. World Wide Web Consortium (2005) (URL: http://www.w3.org/TR/EARL10/)
15. Sidar Foundation. Hera tool Version 1.0 (2003) (URL: http://www.sidar.org/ex_hera/index.php.en)
16. The PHP Group: PHP Hypertext Preprocessor (2006) (URL: http://www.php.net/)
17. Bos, B.: CSS: Cascading Style Sheets. World Wide Web Consortium (2006) (URL: http://www.w3.org/Style/CSS/)
18. Opera Software: Opera Web Browser (2006) (URL: http://www.opera.com/)

Automatic Control of Simple Language in Web Pages

Constantin Jenge, Sven Hartrumpf, Hermann Helbig, and Rainer Osswald

Intelligent Information and Communication Systems (IICS)
Department of Computer Science
FernUniversität in Hagen, Germany
http://pi7.fernuni-hagen.de/

Abstract. The use of simple and easy to understand language is an essential requirement for web documents to be accessible by people with cognitive or reading disabilities. We present an architecture that scores the readability of textual content automatically and also provides hints about possible obstacles to readability in the given text. Our approach is based on a natural language processing framework that supports all levels of linguistic analysis, ranging from the morphological analysis of words to the semantic analysis of sentences and texts.

1 Introduction

The vision of an information society implies that everyone can easily access all the information he or she needs from online resources such as the World Wide Web. The advantage of having all kinds of information available, however, comes along with the challenge to make use of it. In order to keep this challenge as modest as possible, the information should be presented in a way that meets the abilities of the intended target group. In particular, textual information should be formulated in a way that is understandable without problems by those people the information is relevant to. Often the target group is rather heterogenous with respect to reading capabilities. Information related to administrational and governmental issues, for example, affects more or less all members of the society. Another important example are health related topics.

If a piece of textual information is to be accessible by virtually everyone, people with reading impairment have to be taken into consideration as well. This group includes individuals with acquired or developmental language disabilities as well as people with poor reading skills either because of a low education level or because the language in question is not their mother tongue. The need for simple, understandable language in the context of web accessibility has found its expression in the Web Content Accessibility Guidelines (WCAG) proposed by the Web Accessibility Initiative (WAI) of the World Wide Web Consortium (W3C). The respective guideline reads: "Make text content readable and understandable" (see also Section 2.2). Even though an easy to understand language of course benefits all users, the guideline explicitly aims at helping people with reading disabilities.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 207–214, 2006.

Since creating web pages without accessibility barriers is meanwhile being recognized as an important issue for governmental or commercial content providers, there is a growing demand for tools that support the automatic testing of accessibility guidelines. In this paper, we introduce an architecture that supports the automatic evaluation of web pages with respect to readability. The work presented here is part of the EU funded BenToWeb project (Benchmarking Tools and Methods for the Web). One of the explicit aims of this project is to investigate the feasibility of automatic testing of WCAG accessibility guidelines. Although the implemented prototype is at present restricted to the German language, our approach should carry over to other European languages as well.

2 Approaches to Readability Control

2.1 Style Guides, Controlled Languages, and Readability Formulas

There have been several lines of work targeting readability. The most influential are style guides, controlled languages, and readability formulas. Guides to style and composition are probably as old as the art of writing itself. They consist of hands-on advice like "Use short sentences" or "Prefer familiar, everyday vocabulary". Style guides are for the most part based on intuition and common sense. They can be of great help to human authors in writing readable text but often yield criteria that are difficult to control automatically. Meanwhile, there are various initiatives to develop guidelines for authors of official documents, including the promotion of plain language by the U.S. Government.[1]

Controlled languages are applied in technical documentation and in machine translation. The specifications of such languages aim at reducing the complexity of natural languages by constraining vocabulary and syntactic phenomena and by avoiding ambiguity. Controlled languages are thus to be considered subsets of natural languages. A prominent example is ASD[2] Simplified English, which has become the de facto standard for technical documentation in the air and space industries worldwide. Technical writers at Boeing use the Boeing Simplified English Checker (BSEC, [22]) to check their documents for compliance with the language standard. Controlled languages are the only line of work within readability for which automatic compliance checking technologies exist. But we must keep in mind that these approaches are special purpose, limited domain languages and that the corresponding support technologies have to cope only with limited linguistic complexity.

Readability formulas were first developed in the 1920s in North America to determine the adequacy of reading material for classroom use. A prominent example is Rudolf Flesch's Reading Ease Scale [7]. Readability formulas typically exploit superficial statistical text properties like average sentence length, average number of syllables per word, etc. While these formulas have often been

[1] See `http://www.plainlanguage.gov/`.

[2] ASD stands for the AeroSpace and Defence Industries Association of Europe.

criticized as unreliable (e.g. [6]), they where highly influential in teaching and journalism. There have been various proposals for readability formulas suitable to German, either by adapting formulas developed for English or by designing formulas specifically for the German language. The latter type of approach has been pursued by Dickes and Steiwer [5] in the 1970s: Based on 60 German texts whose readability has been scored by the cloze test ([21]), a readability formula was derived via multiple regression from 38 indicators ranging from simple word count to the number of subordinate sentences. For practical purposes, Dickes and Steiwer developed two simplified versions of the formula by the same method, with indicators restricted to those that can be easily calculated by hand or by computer. Since over the last thirty years what can be easily calculated by computers has changed dramatically, and in view of the tremendous progress of automatic natural language processing (NLP), it seems worthwhile to reconsider the readability formula approach, especially when high level semantic indicators are taken into account (see Section 4).

2.2 Readability in the Context of Web Accessibility

Web content is inherently open domain (in the sense that web content may be on any topic conceivable). Authors and readers of web pages need not, and most likely will not, share any common background as, for instance, professional affiliation, ethnic, geographic, or religious origin. Thus, enforcing a language standard of any kind is certainly no option.

As mentioned in the introduction, the Web Content Accessibility Guidelines (WCAG 2.0) require texts to be "readable and understandable" (Guideline 3.1). Interestingly, while the earlier working drafts of the WCAG 2.0 were quite specific with respect to linguistic phenomena and their impact on readability, subsequent versions have seen a progressive dilution of the readability related recommendations. The main reason for removing these recommendations appears to be the fact that many of them only applied to English and European languages, as discussed during the public reviewing of the working drafts. (Another reason may be the lack of methods for automatic testing.) As far as linguistic phenomena are concerned, Guideline 3.1 of the latest WCAG 2.0 working draft[3] is only specific about obligatory explanations for abbreviations and for "specific definitions of words used in an unusual or restricted way, including idioms and jargon." To ensure readability in general, the WCAG draft now refers to the "lower secondary education level" as defined by the International Standard Classification of Education of the UNESCO. Guideline 3.1 essentially says that if a text requires reading ability more advanced than the lower secondary education level then a text summary requiring reading ability less advanced than this education level should be added (or some other type of supplemental content). Put briefly, the W3C Web Content Accessibility Guidelines essentially abstain from listing specific criteria a text must satisfy in order to be linguistically accessible.

[3] Working draft of 23 November 2005.

3 Readability Criteria from a Linguistic Point of View

Since there is no canonical set of criteria for linguistic accessibility, we have developed a catalog (see [16]) based on the following two conditions: All criteria must be psycho-linguistically motivated and must be testable with state-of-the-art natural language processing technology.

We have categorized the criteria according to the level of linguistic description that applies to them. The five levels we distinguish are morphology, lexicology, syntax, semantics, and discourse. The following criteria are a subset of investigated criteria; the selection here includes mainly criteria that are or will be implemented in the DeLite system (see Section 5).[4]

Morphology. Affixation and derivation should result in a predictable outcome. Morphologically complex words, especially when they are infrequent, can impose a considerable burden to the human language processing system [20,19,1]. *Composition** is a highly productive phenomenon in some languages (German is a notorious example). Compounds can become very long and complex, which makes them hard to read. But, what is more, the internal relation between the individual members of the compound may not always be clear. *Abbreviations and acronyms** can unnecessarily obscure the text. They must be explained when they first appear and ideally also in a glossary. In HTML documents, the abbreviation and acronym elements (⟨abbr⟩ and ⟨acronym⟩) must be used.

*Lexicology. Word frequency** plays a crucial role in language understanding. Words that frequently occur are more likely to be familiar to readers. Furthermore, plausibility, i.e. the high probability of a word occurring in a certain context, is central to disambiguation processes [23,10]. *Lexical ambiguity** can irritate and should be avoided if possible. If ambiguity is unavoidable, the context should be such that the intended meaning can be identified easily. Words and expressions that belong to an unusual *register* (e.g. slang, elevated speech) or to *specific domains* should be used with good reason only. Depending on the targeted audience, an explanation or definition may be in place. *Idioms* are phrases where the intended meaning deviates from the literal one. Idioms do not contribute to readability and especially second language users' knowledge of idioms may be limited. *Support verb constructions* (SVC) are phrases like *to make a proposal* versus the simple verb *to propose*. SVCs consist of a verb with an unspecific meaning (*make*, *do*, *go*, etc.) together with a noun phrase or an adjective that contributes the main semantic content of the phrase. For English, SVCs are considered to increase syntactic complexity without gaining expressivity, while they are central to other languages (e.g. Hindi, Farsi). Thus the applicability of a criterion *prefer simple verbs to SVCs* has to be established for each language individually.

[4] The asterisk (*) marks criteria for which tests will be implemented in the first prototype of the DeLite system as our contribution to the BenToWeb project.

*Syntax. Structural ambiguities** come in various flavors. Most prominent are attachment ambiguities (e.g. *The policeman shot the gangster with the gun*). *Structural complexity** can be considered a property of the syntactic representation of a sentence, e.g. the depth of nesting, the length of paths in syntactic representations like dependency graphs and constituent trees [18,17,8,9]. *Sentence length** is also a measure, but admittedly not a very sophisticated one. Most style guides suggest not to exceed 20 words. As such it has proven its worth in the past. But it is also a legitimate indicator for fall back strategies in case linguistic analysis fails. *Clusters* of nouns, adjectives, or adverbs increase the density of information. They should therefore not exceed certain limits. ASD Simplified English does not allow more than three nouns in a row, but noun clusters in English often appear disguised as compounds in German (see above).

*Semantics. Semantic complexity** might be conceived as a metric considering the number of propositions or conceptual entities in a text unit. Whatever the metric, semantic complexity should be constrained. *Reference ambiguity** arises when a pronoun can refer to more than one noun phrase in the foregoing text. It should always be possible to uniquely match a pronoun and its antecedent. *Reference distance** is the distance between a pronoun and its antecedent. The longer the distance, the harder it is to match the two [4,2,9].

Discourse. Discourse coherence has been identified as vitally important for readability. Points to mention here are the use of so called *coherence devices*: sentence connectors like *thus*, *furthermore*, etc. A consistent theme-rheme structure and the proper use of pronouns belong here as well [11,2]. For most of these discourse phenomena, there are no robust analysis techniques. DeLite tries to detect pronouns and noun phrases that have no antecedent although they require one to be understandable.

4 Linguistic Analysis and Readability Indicators

Since we aim at a fully automatic scoring of text readability, all criteria under consideration must be based on strictly operational conditions. To this end, we specify a variety of *numerical indicators* which in turn are defined in terms of the linguistic units and constructs of the text and their features.

Basic indicators encode number information, with length information as a special case, and can apply to all levels of linguistic analysis. Simple examples are the number of word *tokens* in a sentence or a text (i.e., sentence and text length in words), the number of word *types*, and the number of *lemmas* (base forms of words). (Especially for highly inflecting languages like German the latter two numbers can differ considerably.) Examples of basic indicators related to syntactic complexity are the length (in words) of a nominal phrase and the number of embedded constituents (of the same type). Indicators related to semantic complexity include the number of concepts and the number of propositions in a sentence.

Almost all of the above indicators presume linguistic processing ranging from elementary tasks such as lemmatizing to complex tasks such as full semantic analysis. Whereas it is fairly canonical what the lemma of a word is, the syntactic and the semantic representation of natural language presuppose a commitment to some sort of descriptive framework. The natural language processing component of the DeLite system described in Section 5 returns *dependency trees* and *semantic networks* as syntactic and semantic representation, respectively. In our choice of the indicators, care was taken to reduce the dependency on the chosen form of syntactic and semantic representation to a minimum. For example, even if it might be arguable in detail how to count the propositions a concept expressed in a text is involved in, it is not essential that we make use of a semantic network representation (though we think that this kind of representation is preferable to others with respect to cognitive adequacy; see [14]).

At present, we have defined about 50 basic indicators. Several of these indicators are directly connected to possible violations of the readability criteria. If, for instance, the indicator for clause embedding depth exceeds a certain threshold then the sentence in question violates the criterion of syntactic simplicity. The basic indicators are assigned to specific parts of the text, e.g., to a certain sentence or constituent, and can hence be used to locate potential readability problems.

Basic indicators can be combined into *derived* ones in various ways. Standard examples are the arithmetic mean, the median, or the maximum of indicators of the same type such as the average word length in a sentence or the maximal number of prepositional phrases per sentence in a text. In the foregoing definitions of derived indicators, the underlying basic indicators are typically associated with consecutive linguistic entities (e.g. words or sentences). Another option is to combine indicators of different type that are assigned to the same entity. For instance, the various indicators related to syntactic complexity can be combined into a weighted sum which gives us a single (derived) indicator of syntactic complexity per sentence. Obviously, well-known measures like the type-token ratio and other readability formulas (see Section 2.1) can be regarded as derived indicators as well.[5]

5 The DeLite System

The DeLite system implements the method of readability testing by evaluating linguistic indicators as briefly described in the previous section. DeLite will serve a twofold purpose. As an authoring tool, it identifies and highlights problematic text passages in web pages and gives hints towards the nature of the problem. As an evaluation tool, it rates the readability of a web page's text. At present, a prototype is available that works on German web texts.

[5] It is argued in [3] that classical readability formulas are not well suited for web pages because the latter often consist of short passages only – especially in the presence of additional graphical material.

In order to determine the linguistic indicators of the input text, DeLite employs a syntactico-semantic parser [15,12] that provides an analysis of the text at all levels of linguistic description.[6] In particular, the parser returns lemma information for each word, a syntactic dependency tree and a semantic network for each sentence, and a list of possible antecedents for every noun phrase. DeLite then aggregates this information into an annotation structure for the text in question and assigns the basic and derived indicators to the respective components of the structure. On the basis of this data structure, DeLite can report on readability violations and scores via a graphical user interface or a structured report form.

6 Future Work

Plans for future work encompass an thorough evaluation of the system, which includes a validation of the criteria catalog and an extension to the English language. With respect to calculating overall readability scores we are planning to set up a machine learning scenario for adjusting indicator parameters automatically by applying our system to texts that are pre-classified (annotated) with respect to certain readability aspects. Possible corpus sources are, for instance, news and other non-fiction texts especially written for children[7] and text variants resulting from professional readability revisions of texts. Another topic of future research especially virulent in the context of the web is concerned with the automatic adaption of our system to specific domains (e.g. those providing governmental or health content), where specific expressions can be far more frequent than usual without affecting readability.

Acknowledgments

Constantin Jenge has been supported by the EU project *Benchmarking Tools and Methods for the Web* (BenToWeb, FP6-004275).

References

1. Harald Clahsen. Linguistic perspectives on morphological processing. Unpublished, 2004.
2. Charles Clifton and Fernanda Ferreira. Discourse structure and anaphora: Some experimental results. In Max Coltheart, editor, *Attention and Performance*, volume 12, pages 777–793. Erlbaum, Hillsdale, New Jersey, 1987.
3. Kevyn Collins-Thompson and Jamie Callan. Predicting reading difficulty with statistical language models. *Journal of the American Association for Information Sciences and Technology*, 56(13):1448–1462, 2005.

[6] The parser makes use of a large semantically-based computational lexicon[13].

[7] Fairy tales, though mostly written for children, are of limited use because they often contain metaphors and outdated words and the constituent order often follows rhythmic goals and is thus unusual.

4. Meredyth Daneman and Patricia A. Carpenter. Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, 19:450–466, 1980.
5. Paul Dickes and Laure Steiwer. Ausarbeitung von Lesbarkeitsformeln für die deutsche Sprache. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 9:20–28, 1977.
6. Thomas M. Duffy. Readability formulas: What's the use? In Thomas M. Duffy and Robert Waller, editors, *Designing Usable Texts*, pages 113–143. Academic Press, Orlando, Florida, 1985.
7. Rudolf Flesch. A new readability yardstick. *Journal of Applied Psychology*, 32:221–233, 1948.
8. Edward A. F. Gibson. *A Computational Theory of Human Linguistic Processing: Memory Limitations and Processing Breakdown*. PhD thesis, Carnegie Mellon University, Pittsburgh, 1991. Unpublished doctoral dissertation.
9. Edward A. F. Gibson. Linguistic complexity: Locality of syntactic dependencies. *Cognition*, 68:1–76, 1998.
10. Edward A. F. Gibson and Neil Pearlmutter. Constraints on sentence processing. *Trends in Cognitive Sciences*, 2(7):262–268, 1998.
11. Michael A. K. Halliday and Ruqaiya Hassan. *Coherence in English*. Longman, London, 1976.
12. Sven Hartrumpf. *Hybrid Disambiguation in Natural Language Analysis*. Der Andere Verlag, Osnabrück, Germany, 2003.
13. Sven Hartrumpf, Hermann Helbig, and Rainer Osswald. The semantically based computer lexicon HaGenLex – Structure and technological environment. *Traitement automatique des langues*, 44(2):81–105, 2003.
14. Hermann Helbig. *Knowledge Representation and the Semantics of Natural Language*. Springer, Berlin, 2006.
15. Hermann Helbig and Sven Hartrumpf. Word class functions for syntactic-semantic analysis. In *Proceedings of the 2nd International Conference on Recent Advances in Natural Language Processing*, pages 312–317, Tzigov Chark, Bulgaria, 1997.
16. Constantin Jenge. BenToWeb deliverable D6.1. Technical report, FernUniversität in Hagen, 2005.
17. Susan Kemper. Life-span changes in syntactic complexity. *Journal of Gerontology*, 42:323–328, 1987.
18. John Kimball. Seven principles of surface structure parsing in natural language. *Cognition*, 2:15–47, 1973.
19. Thomas F. Münte, Tessa Say, Harald Clahsen, Kolja Schiltz, and Marta Kutas. Decomposition of morphologically complex words in English. *Cognitive Brain Research*, 7:241–253, 1999.
20. Robert Schreuder, Marjon Grendel, Nanda Poulisse, Ardi Roelofs, and Marlies van de Voort. Lexical processing, morphological complexity and reading. In David A. Balota, Giovanni B. Flores d'Arcais, and Keith Rayner, editors, *Comprehension Processes in Reading*, pages 125–141. Erlbaum, Hillsdale, New Jersey, 1990.
21. Wilson Taylor. Cloze procedure: A new tool for measuring readability. *Journalism Quarterly*, 30:415–433, 1953.
22. Richard H. Wojcik, James E. Hoard, and Katherina Holzhauser. The Boeing Simplified English Checker. In *Proceedings of the International Conference Human Machine Interaction and Artificial Intelligence in Aeronautics and Space*, pages 43–57, Toulouse, 1990.
23. Rolf A. Zwaan and Gabriel A. Radvansky. Situation models in language comprehension and memory. *Psychological Bulletin*, 2(123):162–185, 1998.

Test Case Management Tools for Accessibility Testing

Sandor Herramhof[1], Helen Petrie[2], Christophe Strobbe[3], Evangelos Vlachogiannis[4], Kurt Weimann[5], Gerhard Weber[5], and Carlos A. Velasco[6]

[1] University of Linz, Institut Integriert Studieren, Altenbergerstr. 69, 4040 Linz, Austria
sandor.herramhof@jku.at
[2] University of York, Department of Computer Science, Heslington, York, YO10 5DD, UK
petrie@cs.york.ac.uk
[3] Katholieke Universiteit Leuven, Dept. Electrical Engineering - SCD, Kasteelpark Arenberg 10, 3001 Heverlee, Belgium
christophe.strobbe@esat.kuleuven.be
[4] University of the Aegean, Department of Product and Systems Design, Ermoupolis, Syros, GR83200, Greece
evlach@aegean.gr
[5] University of Kiel - Multimedia Campus Kiel, Theodor Heuss Ring 140, 24143 Kiel, Germany
{k.weimann, g.weber}@mmc-kiel.com
[6] Fraunhofer-Institut fuer Angewandte Informationstechnik FIT, Schloss Birlinghoven, D53757 Sankt Augustin, Germany
carlos.velasco@fit.fraunhofer.de

Abstract. Two tools are presented which support test case management for accessibility test suites. Creating test suites for the Web Content Accessibility Guidelines 2.0 is one major objective of the EU-funded project BenToWeb[1]. Parsifal is a desktop application which easily allows editing test description files. Test description files compose an XML layer containing descriptive information about the particular test cases. Amfortas is a web application which allows controlled evaluation of the test suites by users. Controlled in that sense means, that Amfortas not only stores the evaluation results, but also is aware of the physical and technical condition of the evaluator.

1 Introduction

According to the AMFORTAS methodology [1] a test case consists of test material (e.g. an HTML file) and a test case description. The Test Case Description Language TCDL [2] developed within the scope of BenToWeb, not only states whether a test case passes or fails. For user evaluations it additionally provides guidance on the combination of assistive technologies/user agents/devices a test

[1] http://www.bentoweb.org

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 215–222, 2006.

case should be tested with. The creation of test cases (a complete set matching each of the requirements of a high-level specification like WCAG) is costly and work-intensive, but it has the advantage that false negatives are much easier to find than in the evaluation reports of real websites [3].

Preparation, editing and management of user, expert or manual tests can be improved if appropriate software tools are available. Two tools have been developed to support the editing process of the test case description files and to manage the test case evaluation procedure.

2 Parsifal – Test Case Editor

Parsifal is a graphical Test Case editor for editing user test descriptions as defined by TCDL1.1 specification [2]. Test case description files are written in XML. Since work with XML documents in text editors is not very comfortable and error-prone a graphical editor was implemented to ease editing XML test case description files. Parsifal is implemented in C# using .Net Framework 1.1. A setup routine for easy installation and runtime configuration is provided for Parsifal. In Parsifal parallel work on test case descriptions is guaranteed by using a Concurrent Versions System (CVS), for managing multiple submits from different partners at a time. The Annotation facility of CVS is used to comment changes made on test cases.

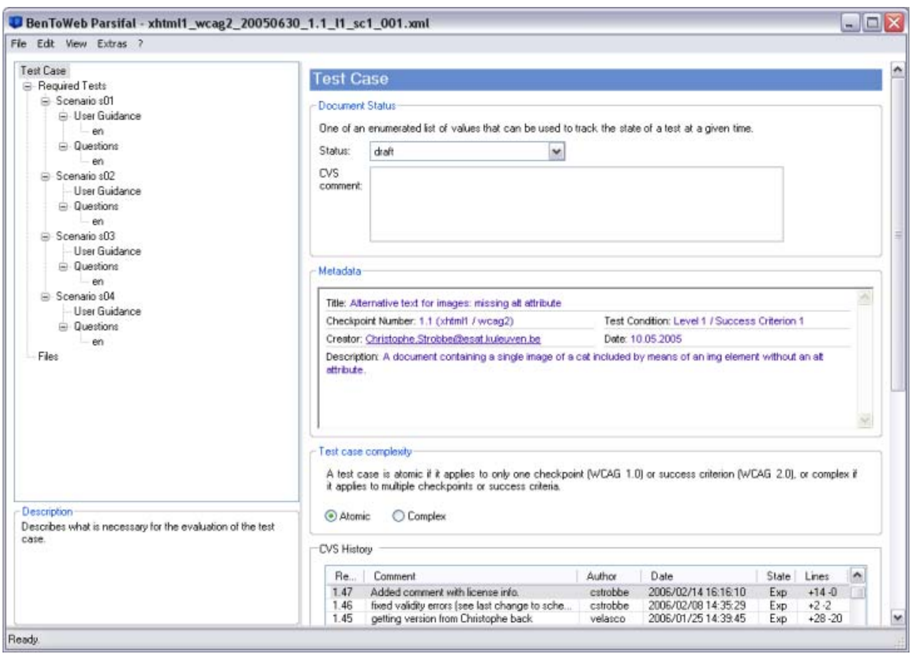


Fig. 1. Status information about a test case description file

Parsifal consists of three main components. A graphical user interface component, which describes the graphical editor. A CVS component, which deals with CVS file synchronization between editor and CVS server. And a serializer component, which is responsible for reading, writing, previewing and printing test case documents. Standard requirements for editing, saving and printing test case descriptions have been implemented in Parsifal. Moreover features like document preview and CVS debugging have been added for internal quality assurance. Parsifal's graphical interface consists of three main panes (see Fig. 1.). Document structure and navigation tree on the top-left. A short description about the meaning of the tree entry is on the lower left, the main pane (content pane) is on the right. The top level tree node is called 'Test Case' and gives the author a detailed summary to which WCAG guideline the test applies. In addition test case meta-data includes information about the WCAG guideline the and success criteria the test covers. This information is relevant for test case authors preparing user tests.

In 'Document Status' field authors can change the document's status (e.g. draft version) and can add a CVS comment to their changes made. Document history and versions are also traceable on this pane. The section labeled 'Files' provides access to relevant test files. Test files are presented in the user panel. On basis of a test file the authors can create test scenarios. Therefore authors can preview test files and read the tests purpose. The test case purpose describes what the test file should test and what the expected test result is. The section 'Required Tests' hosts scenario child nodes describing user tests. Tests can be marked as automatic, end user, experts, and one expert test by the test case authors. A user test description is only necessary if a test is not marked as automatic test only. Authors can remove and add scenario child nodes as needed. A scenario (see Fig. 2.) describes user groups to run the test with corresponding questions and user guidance. A user is categorized by his disability/disabilities and his experience level in usage of assistive technologies, user agents and devices. For all three experience levels detailed information about each type (e.g. browser, screen reader), exact product name, product version, and experience level can be specified. Experience levels range from not experienced to very experienced (1-5). One scenario node includes user guidance and questions subsections. User guidance and questions will be presented to users in multiple languages in Amfortas. Therefore localization of all texts is possible in Parsifal. One user guidance entry describes necessary prerequisites a user must undertake to accomplish a test. One sample instruction could ask the user to switch on his screen reader. All user guidance and question texts are described in a subset of XHTML1.0. Therefore a basic HTML editor component eases authoring user guidance and question text entries. Parsifal supports questions of type yes/no, yes/no/open, multiple choice, and likert scale. Each question type supports different settings like likert scale and multiple choice questions need at least labels and values being specified. Open questions need information about layout and spacing to be presented correctly to user panel by Amfortas. As mentioned before all question texts must be translated into different languages, this happens by adding child nodes to the Questions entry with corresponding language specific labels.

On the right hand side question corresponding input can be edited. In this case a yes/no/open question at least needs a question text. A mandatory 'Other Options' entry is used if e.g. yes/no is not an appropriate answer to users. In this example the user was asked to navigate through the page using the tab key and if he got the welcome message. Most of the users will answer with 'yes' or 'no' but those who have e.g. technical difficulties would like to answer more precisely. Therefore 'Other Options' asks for the reason why the user was not able to get the information propagated.

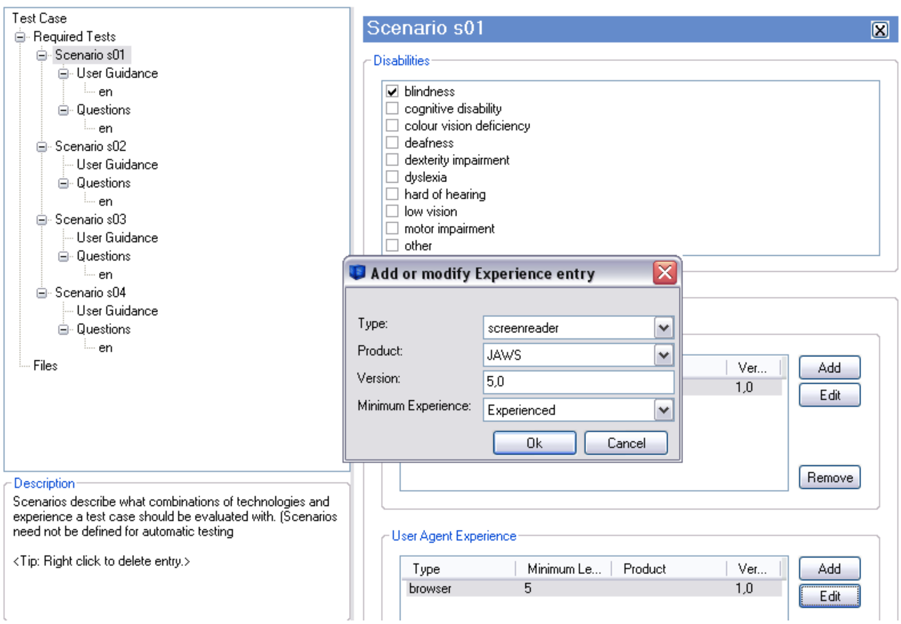


Fig. 2. Scenario pane in Parsifal

3 Amfortas – Test Case Evaluation Framework

The framework is a Java-based web application for evaluating accessibility test suites. It covers the whole management process from creating and handling user and testing profiles, to storing the test results in a database. Usually user evaluation processes are monitored by an agent, who in a first step presents a task and then gathers information about the process and the result by asking relevant questions. Obviously, this makes testing fairly expensive. The evaluation framework should ease user testing procedures in a way that evaluators can proceed testings by themselves without any human support.

3.1 Process Work-Flow

The evaluation task starts with a recruitment procedure. The recruit is guided through a series of questions in order to gather information about his personal constitution (disability, age range, internet experience) and technical equipment (assisitive technology, browser, device used to access web content). The answers of the recruitment procedure determine one testing profile. The evaluator should be able to set-up additional profiles, if for some reason the equipment changes or he uses more than one set of equipment to access web content.

The administrator can view the status and profile of the registered participants. All participants with adequate profile will be granted access to the evaluation framework by activating their accounts. The users then are able to access the log-in area of the web-application, but cant start testing unless the test profile is admitted to a particular test suite.

The testing process starts by activating the corresponding link in the web application. The mapping algorithm first looks up the database, selects profile information for the actual user and tries to match it with the TCDL description files. The matches are stored in a pooling table, in blocks of 20 test cases (called test run). Testing can be repeated as long as there are matching test cases available. The evaluators are expected to have at least moderate English skills. As all test cases are in English, the log-in area of the web-application is also kept in English.

The framework guides the evaluator through all the allocated test cases. A test case is finished when the user answers the question about the test case. The answer is stored in the database.

After evaluating the test suite the data is extracted from the database for later analysis.

3.2 System Architecture

Amfortas is built on top of the XML publishing framework Cocoon[2]. The core of Cocoon's object-oriented architecture is based on the Apache Avalon project. The overall architectural view of Amfortas consists of three components: a Java web server containing the application, a MySQL database and a resource containing the test files and test description files. Usually, the files are provided via a web interface, but any other providing mechanism, e.g. CVS, would also be appropriate.

Due to Cocoon's internal architecture, the evaluation framework is composed out of 3rd party components, own components, Javascript files, XML files and certain additional resources.

Database Layer: Instead of creating a custom persistency layer, we decided to use Hibernate[3]. Hibernate not only provides a powerful and easy-to-use object-relational bridge for Java applications, but also offers a rich query language to retrieve objects from the database.

---

[2] http://cocoon.apache.org/

[3] http://www.hibernate.org

The evaluation framework uses 41 tables to store persistent data. Persistent data is data needed to build up the application view, data to accomplish the mapping procedure, data to conduct application management procedures and data which composes the evaluation result. The database model is straightforward: it is actually a normalized view on the users personal condition and technical equipment. Amfortas stores the mapping-related data for assisitive technologies, user agents, devices and disability in different tables, which in the end are consolidated in the table `test_profiles`. One entity set in `test_profile` determines one test profile.

Presentation Layer: The initial version of the user evaluation framework presents a very simple and intuitive user interface, as it is going to be accessed by users with a huge variety of interaction requirements. For later extension of functionality a clear separation of content and presentation is needed.

Amfortas' content has been completely authored in XML reusable entity documents. This process actually involves three sitemap components. If there is no need for aggregation, a Cocoon generator simply loads XML from the file system or web resource and generates SAX events which are handled by consecutive XSLT transformers and finally a serializer (i.e. HTML for Browsers). In most cases, an Cocoon aggregator is required which offers additional functionality by aggregating more than one XML files (i.e. Header, Content and Footer) inside a root element.

Amfortas' public access area is implemented in the languages English and Dutch, as the potential evaluators are recruited in England and Belgium. Cocoon offers the i18n-Transformer component to implement internationalization features. Language-dependent text is stored in an XML file and referenced by the application through a unique key.

CForms - Forms are important for interaction but at the same time raise a lot of accessibility issues. This is mainly due to the need for direct and responsive interaction, which is usually implemented with client-side technologies - in most cases Javascript - which may cause serious accessibility barriers. These problems have been already addressed by W3C, which proposes the next generation of web forms named XForms. XForms seems ideal for Amfortas' forms implementation, but it's not applicable as most user agents have not yet implemented this technology. A good alternative that merits goods from both current and future world are Cocoon forms (CForms). CForms are XML forms that introduce the separation of the model (form model) and instance (form template) of the form that can be implemented separately. The so-called form widgets can be developed and include their own server-side validation. In Amfortas, the form instances are controlled with Cocoon flow. A further advantage of this approach is the ability to move to XForms by simply applying XSL transformation.

Application Logic: While 'action components' have been the dominant method to encapsulate application logic in Cocoon, this position has been taken over by the *Control Flow*. A flow script is implemented in Javascript notation.

A considerable part of application logic, like the recruitment process, the application administration or the saving routine of the evaluation results is implemented using flow scripts. Application logic which directly influences the view of web application is implemented using the Cocoon JX-Transformer. Higher-level logic, such as the mapping procedure, is implemented in Java classes.

Profile Mapping. Only test cases marked with status 'accepted for end user evaluation' are pooled. Each user request for new test cases triggers the mapping algorithm. Mapping involves comparing TCDL disabilities and experience elements (i.e. user agents, assistive technologies and devices) with the user's test profile stored in Amfortas' database. The mapping algorithm first filters out the test cases that are 'done' and those that the user has already evaluated. For a test case to be allocated, the following conditions needs to be satisfied. For disabilities, if in TCDL there is a disability, the test profile needs to have at least one of them. For user agents, assistive technologies and devices, the test profile must have all the types that appear in TCDL file. If the TCDL specifies a product as well, the profiles need to have at least one of the specified products for each type. Further, if minimum level and product version are also specified these need to be equal or less than the profile's one. Finally, for better matching, a complementary grading mechanism is involved that enables a better selection after sorting by grade and getting the required number designated by the test suite configuration.

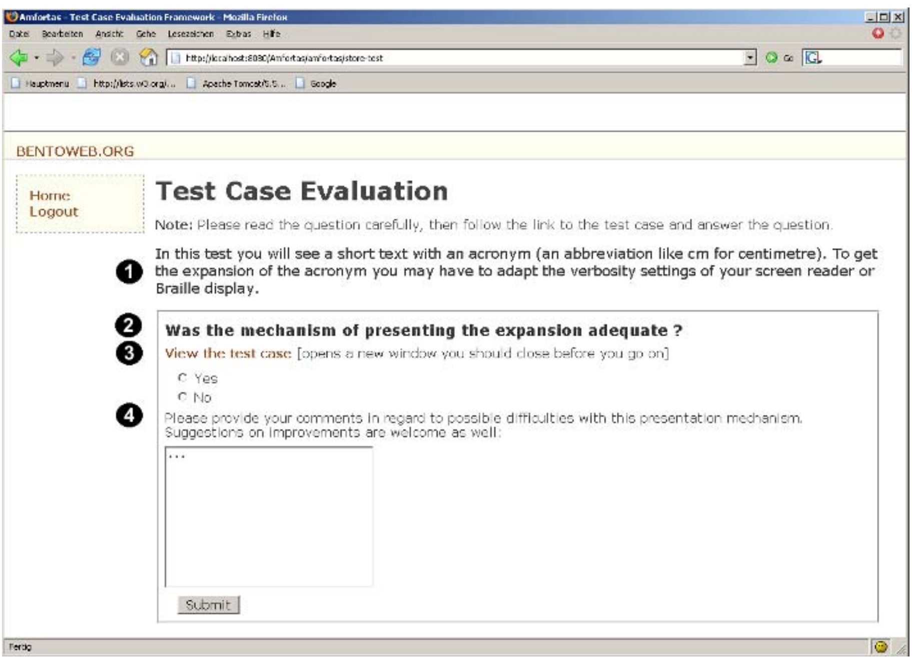


Fig. 3. Amfortas test case evaluation

Test Presentation. The test case evaluation is a cyclic process which can be invoked as long as test cases can be allocated for the actual evaluator. The whole test presentation is created within the Cocoon sitemap, by passing a number of parameters from the Cocoon flow, for example, the URI of the actual test case description file, the URI to the test file, and the scenario. The default sitemap generator is used to fetch the test description file from the web resource and page header, footer and navigation from the local file system. The standard cocoon aggregator bundles these XML trees to a single XML tree which is handed over to XSLT transformation and finally serialization. The result is an XHTML page with user guidance information, a link to the test file and a question with corresponding answer type (see Fig. 3.).

3.3 Application View

Fig. 3. shows the user interface for the evaluation process. The 'user guidance' section (1) requests the evaluator to adjust special settings or behave in a certain manner in order to complete the test. The question (2) is presented before the link to the test file (3), to give a first idea what to mention when evaluating the test file. Finally the answer section (4) presents one of the answer categories to be replied. On submit, the answer with references to the accomplished test is stored.

Acknowledgements

This work has been undertaken in the framework of the project BenToWeb IST-2-004275-STP funded by the IST Programme of the European Commission.

References

1. Petri, H., King, N., Gappa, H., Nordbrock, G., Velasco, C.: Large scale evaluation of the accessibility of Web features. Computers Helping People with Special Needs. 10th International Conference, ICCHP 2006, Linz, Austria, 12-14 July 2006, Proceedings.
2. Strobbe, C., Herramhof, S., Vlachogiannis, E., Velasco, .C.: Test Case Description Language (TCDL) - Test Case Metadata for Conformance Evaluation. Computers Helping People with Special Needs. 10th International Conference, ICCHP 2006, Linz, Austria, 12-14 July 2006, Proceedings.
3. Brajnik, G.: Comparing accessibility evaluation tools: A method for tool effectiveness. Universal Access in the Information Society. Springer Verlag (2004), pp. 252-263.

Testing Web Navigation for All: An Agent-Based Approach

Martin González, Marcos González, Cristóbal Rivera,
Ignacio Pintado, and Agueda Vidau

The Human Communication and Interaction Research Group, University of Oviedo,
Department of Computer Science, Calvo Sotelo s/n, 33007, Oviedo, Spain
martin@uniovi.es, {marcosgonzalezgallego, nacho_x8}@yahoo.es,
xtoball@gmail.com, avidau@gmail.com

Abstract. Laboratory navigability testing is a powerful technique to obtain a picture of the user's mental representation of the navigation model of a web site. However, bringing volunteers to the laboratory to test our prototypes is quite expensive and even impossible for certain users with specific interaction needs. Since the tests are performed using equipments different from those employed by the real user, the impact that the computer performance has on navigation is missed. In this paper we propose a remote testing approach, performing navigability testing in the user's home, employing special silent data-gathering software agents, which are able to measure the user accuracy when performing navigation tasks.

1 Introduction

Web navigation is as process where decisions must be made continually, regarding strategies for reaching the goal and determining whether the goal has been reached [1]. People need to determine what is in their environment and how to find their way around [2]. As navigation strategies are based on topological considerations of the knowledge structure [1], orientation involves the alignment of the user's mental representation of the navigation model with the designer's mental representation.

Since the user sometimes follow a plan and sometimes respond to the environment [2] it does not seem to exist a rule of thumb to achieve the goal of a perfect alignment between both mental representations. As a result, testers must expose their web products to real users performing periodic usability and navigability testing [3] before releasing the final product.

Usability testing offers a powerful and accurate tool to get valuable information that helps designers to analyze the weak points of their site: difficulty of use, misconceptions in the use of the interface, misunderstanding in the classification of contents and much many others [4].

However the data gathering process used by the classic usability testing approach has also an important number of drawbacks, making it inaccurate and expensive when using it at the navigability testing stage of this kind of projects. The main drawbacks are cited next:

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 223 – 229, 2006.

- *Expensive and Difficult Process.* Navigability testing requires accuracy and permanent observation of each volunteer participating in the navigability tests. Observing times are long and it is required at least an observer per volunteer.
- *Destruction of Spontaneous Behaviour.* Volunteers know that they are under observation. The presence of a video cameras required to record their behaviour perceptively modifies people's attitude. This knowledge has a deep impact on navigability, as volunteers tend to adapt their own internal navigational metaphors to the navigational model provided by the researchers conducting the experiment.
- *Difficulties identifying new Kinds of Users.* The testers choose volunteers among people with the same cultural background of the intended target user. This might not be the case of the final user, as any web product will be used by a huge variety of different users.
- *Users with special interaction requirements usually do not participate in the process.* Certain kind of users who need special interaction requirements (such as blind users or people with limited movement) usually do not participate in those testing due to the cost of the special hardware required. Although this equipment is usually available at the user's home, the cost of installing the data gathering equipment in this environment is really high.
- *Navigational Environment is not considered.* Navigability testing takes place behind the walls of a laboratory, using computers with different performance that those employed by the real users. Since several authors reported that navigation is affected by the speed of the equipment [5, 6, 7, 8 y 9] a really important part of the whole picture is missed in this kind of tests.

2 Testing at Home: The Remote Testing Approach

In order to avoid the problems commented, we have proposed the Remote Navigability Testing [10], replacing human observers by software data-gathering agents. The replacement of human observers by software agents is possible if we consider the simple nature of the information to be gathered (nodes and link used as landmarks, reaction times, accuracy, etc.). An automatic data gathering process means a cheaper testing process is cheaper, since the high costs derived from navigability testing are generally caused by hiring a laboratory and human resources for long times of observation.

In the classic usability testing approach, the user visits the laboratory to test software. In the remote testing approach, the software visits the user's home in order to be tested. The prototypes are uploaded to a testing server to be used at home by any volunteer participating in the tests. Since the volunteers are real visitors of the target web site, this approach encourages the detection of new kind of users and allows the participation of users with special interaction requirements with no cost for the researchers (the user is using his own equipment at home).

Since the tests are performed in the user's home, external factors such as stress or the sense of being under observation disappears, obtaining more accurate results. Users can use their own interaction equipment performing navigation tasks in a natural (spontaneous) way, being possible to obtain important data about how the special features of the computing environment might affect navigation for each user.

This technique facilitates experimentation, focusing testing in only one-task (navigation) using special tools to support remote navigability testing. These tools must be able to capture information about the user navigation session, and send it to the testers for an offline analysis.

3 ANTS (Automatic Navigability Testing System)

The remote testing technique requires the use of tools to gather information silently. WAMMI [11] and WebCAT (Web Category Analysis Tool) [12] are evaluation tools based on questionnaires that the users fill out, providing a measure of how easy to use they think a web site is. Although they are able to obtain important information about the user navigation model, they aren't advisable to perform navigability testing as they are unable to record spontaneous navigational behaviour. Silent data-gathering tools (whose presence is unnoticed by the user) should be used instead.

WebVIP (Web Variable Instrumented Program) [13] and EISUR (ErgoLight's Interactive Site Usability Reports) [14] are good examples of this kind of tools. At design stage, WebVIP allows the selection of the types of user events to log. These choices are converted to JavaScript code, which is then inserted inside each page of the target web site. When the user interacts with the web documents, the selected events are captured and recorded in session-specific log files, which can be analysed offline. EISUR works slightly different, analysing web server logs in order to obtain metrics of the usability of the navigation models, as well as a site rating report.

Since those tools obtain and classify the information included into the log archive of the web server, important information about the impact of the user computational environment is missing. Additionally, it is also impossible to obtain information about the user interaction requirements with this kind of tools. For instance, it is not possible to obtain the mouse motion accuracy of the user whether the user prefers a text-based navigation or a mouse-based navigation.

Due to the lack this kind of features in the state of the art of the data-gathering technology, we designed a new tool based on independent agents, which are able to obtain data about specific interaction requirements. This tool is called ANTS (Automatic Navigability Testing System).

3.1 ANTS Inside

The design of ANTS is based on a client-server paradigm, using a design metaphor inspired in colonies of ants. The ants (agents) depart from their anthill (a server) looking for food (the information) in every picnic available (navigation session). Once the ants get their food, they come back to the anthill carefully storing the information in one of the anthill's data warehouses (the data base).

As in real life, where there are different types of ants specialized in performing specific tasks (warriors, workers, etc.) Our system uses a large collection of different ants, which were specially designed for the purpose of gathering specific kind of data. For instance, our system uses text agents to determine user's typing skills, motion agents to measure the user's motion accuracy and so on.

In order to perform a remote usability test, the agents must be included in each web document under testing. Depending on the kind of data to be retrieved, more than one kind of ant should be included in each document. Whenever the researcher tests a small number of static web pages in a hired web server, the ants may be included manually, inserting small pieces of code in each page. However, in wider scenario, whenever the researcher uses his own web server to test a huge amount of static and/or dynamic web pages, the ants are automatically inserted by our ant-proxy tool.

This tool works as a proxy between the web server containing the web project under testing and the users. When the ant-proxy receives the user requests for a specific web page, it will redirect it to the real web server which will return the requested page to the ant-proxy. Prior to return it to the user, the ant-proxy inserts the required ants in the page automatically.

This tool is fully configurable by the tester, using different Http ports for different tasks. For instance, it is possible to use a working ants-free version of the web project and its equivalent testing version (including ants) in the same machine and at the same time. For example, let suppose that a researcher assigns the port 4040 to the ant-proxy in a site called `www.myproject.com`. In this case, the URL of the ants-free version of this site would be `www.myproject.com` (any request on port 8080 will be processed by the web server) while the testing version (including ants inside) would be `www.myproject.com:4040`.

3.2 Data Retrieved with Ants

Whenever a testing web page is retrieved by the user, the ant is downloaded from the anthill (the ANTS server) to his computing device. Once the ant has arrived to the user machine, it establishes a communication channel with the anthill, which is used to send the data collected. An ant-assistant object located in the server-side (see figure 1) maintains the communication channel alive during the interactive dialog's cycle of life, collecting the information sent by the ant, classifying it, and sending it to the warehouse object commissioned to store it in the proper user model.

Any kind of ant is able to gather information about the exact amount of time employed by the user in each web document, including time of arrival, time of departure and his next destination. The agents only measure the time spent by the user when performing navigation task, so the time is not measured when the user stops navigation to perform any other kind of task.

This information is used to create an abstract navigation model including series of nodes linked to each other, where every node represents a web document. The navigation model, is used to determine the landmarks, routes and mental models [1] employed by the volunteers when they navigate thorough the web document. Combining the information gathered from all the volunteers it is possible to obtain the *hot* and *cold* zones of the web site, that is, the most and less visited page clusters. This information is also used to detected popular links and links that are never used.

It is also possible to determine the user's navigation speed comparing his navigation behavior with the average navigation speed in the site. This information is also employed to determine the user's expertise degree [15].

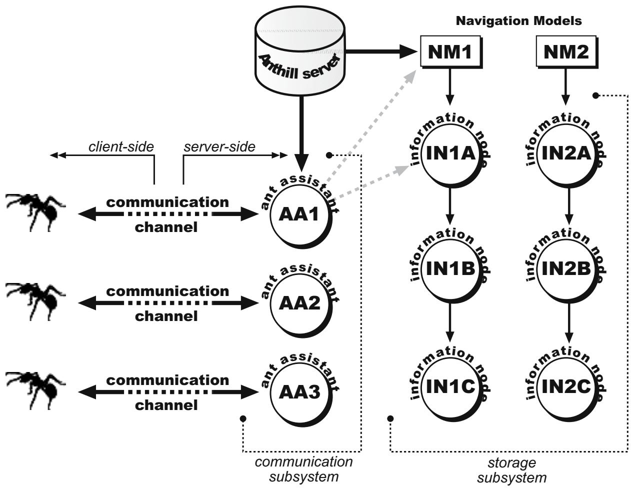


Fig. 1. ANTS navigation model

Specialized agents increase the amount of gathered information or complement it for specific usability tests. When using standard ants we get valuable information about the navigation behavior of a huge amount of users but not for all. For instance, standard ants can get information about the links employed by the user when performing a mouse-based navigation but no information is retrieved when using a key-based navigation.

Visually handicapped users and any other users with similar special interaction requirements need to perform key-based navigation. In this case, the *tab* key is used to select the required link moving the focus from one link to the next one. In web documents designed according to the W3C's WAI Accessibility Level AAA [16], it is also possible to access to specific links using Access Keys (also known as keyboard shortcut) represented by the *accesskey* HTML tag.

Obtaining information about the key pressed by the users along their navigation session is crucial in this kind of tests, so the standard ants are complemented by text ants which are able to record any single key pressed by the user, storing this information in the ant hill. This data is later used by the researchers in order to reproduce the navigation session for an offline analysis.

Mouse-based and key-based navigation gathered information can be analyzed separately in order to determine different navigation speeds, level of expertise and interaction requirements in both approaches.

ANTS also includes other agents specialized in gathering specific information to obtain a better view of the interaction requirements of the users. For instance, the motion-ant is able to obtain a mouse motion accuracy rating for each user of the project. This information is quite valuable in order to design the visual appearance of the web. For instance, if the average users are elderly people or any other kind of users with low motion accuracy, the size of the target objects in mouse pointing tasks (links, buttons, etc.) should be increased accordingly in order to make this task easier.

In order to obtain this score, whenever the user moves the pointer in the screen the ant analyzes its trajectory according to Fitt's Law [17 pp. 4]. This law predicts the time required for reaching a target when a finger or pointer is moved from a place to another.

4 Conclusions

Remote testing clearly increases the quality and productivity of the testing stage of a hypermedia development process. This approach eliminates irrelevant noise (external factors such as nervousness or confusion), captures additional information (influence of the environment on navigation), gets accurate and reliable information automatically and makes testing cheaper a more efficient.

When using automatic remote testing tools such ANTS, data flows freely from the origin to the server storage system, where it can be analysed off-line. As navigation takes place in the own user-computing environment, there is no need to assign expensive laboratory resources for testing. Testers are free from the boring task of capturing data, so they can focus their efforts in analysing the results obtained and improving the quality of the navigation maps. As there is no need to assign human resources for conducting user observation, the whole process of usability testing is cheaper.

The use of agents is much more powerful that a deep analysis of web server's logs in most of the cases. For instance, the agents used by ANTS are able to detect text-based navigation approaches usually performed by visually handicapped users and are also able to measure the mouse motion skills of the users.

We want to remark however, that remote testing does not pretend to be a substitution of classic usability testing. Our intention is to optimise usability testing for certain special tasks. Remote testing is able to detect actions but it is unable to observe how those actions are performed. For example, with remote testing it is quite easy to get the number of clicks on a button during a navigation session, but it is impossible to know the difficulties that the user could have when he/she performed the click. That is a task where classic usability testing seems to be unbeatable.

References

1. Jul, S., Furnas, G. W.: Navigation in Electronic Worlds: a CHI'97 Workshop. SIGCHI Bulletin Vol. 29, No. 4, October 1997, ACM Press. (1997)
2. Nielsen, J: Hypertext and Hypermedia. Academic Press, Cambridge. ISBN 0-12-5118410-7. (1993)
3. Krug, S.: Don't make me think: Don't Make Me Think! A Common Sense Approach to Web Usability. New Riders. ISBN: 0789723107. (2000)
4. Nielsen, J: A Mathematical Model of the Finding of Usability Problems. Proceedings of INTERCHI 1993: Human Factors in Computer Systems. Amsterdam
5. Johnson, C.: What's the Web Worth? The Impact of Retrieval Delays on the Value of Distributed Information. Proceedings of the Time and the Web Workshop, Staffordshire University, 19Th June 1997
6. Johnson, C., Gray, P.: Temporal Aspects of Usability: Assessing the Impact of Time on User Interface Design. SIGCHI Bulletin Vol. 28, No. 2, April 1996, ACM Press

7. Thomas, R. C.: Temporal Aspects of Usability: Long-term Variation in User Actions. SIGCHI Bulletin Vol. 28, No. 2, April 1996, ACM Press
8. O'Donnel, P., Draper, S. W.: Temporal Aspects of Usability: How Machine Delays Change User Strategies. SIGCHI Bulletin Vol. 28, No. 2, April 1996, ACM Press
9. Byrne, A., Picking, R.: Is time out to be the big issue? Proceedings of theTime and the Web Workshop, Staffordshire University, 19Th June 1997
10. Gonzalez M., Automatic Data-Gathering Agents for Remote Navigability Testing. Pages 78-85. IEEE Software, November 2002
11. Nomos Management (Sweden); Kirakowski, J., WAMMI: Web Usability Questionnaire. http://www.nomos.se/wammi/index.html, 2005
12. VUG (Visualization and Usability Group) WebCAT: Web Category Analysis Tool. http://zing.ncsl.nist.gov/WebTools/WebCAT/overview.html, 2002
13. VUG (Visualization and Usability Group) WebVIP: Web Variable Instrumented Program. http://zing.ncsl.nist.gov/WebTools/WebVIP/overview.html, 2002.
14. Ergolight: Usability software: ErgoLight Interactive Site Usability Reports http://www.ergolight-sw.com, 2006
15. Karn S., Perry T. J., Krolczyk, M., Testing for Power Usability: a CHI 97 Workshop. ACM SIGCHI Bulletin Vol . 29, No. 4, October 1997
16. W3C: Web Content Accessibility Guidelines 2.0. W3C Working Draft 23 November 2005. http://www.w3.org/TR/WCAG20
17. Newell, A. Unified Theories of Cognition, Harvard University Press. ISBN 0-674-92101-1 (1990)

University of Illinois Tools and Techniques for Functional Web Accessibility

Hadi Bargi Rangin

Illinois Center for Instructional Technology Accessibility (iCITA)
Division of Disability Resources and Educational Services
University of Illinois at Urbana/Champaign
1207 S. Oak Street, Champaign, IL 61820

Abstract. Functional web accessibility goes beyond complying with the technical requirements of Section 508 or W3C Web Content Accessibility Guidelines 1.0. Current web accessibility policies and practices favor an "accessible repair" approach to web accessibility, which lead to resources that might meet the technical requirements of accessibility guidelines, but are still not functionally usable by people with disabilities. The University of Illinois at Urbana/Champaign (UIUC) has developed a set of HTML best practices and accessibility management and visualization tools to improve the design and verification of the functional accessibility of web resources. The goal of these practices and tools is to support developers and administrators in creating and verifying the functional accessibility of their web resources. The practices encourage developers to use forward looking web design that improves the accessibility of web resources to everyone, including people with disabilities, by making web resources more adaptable to a wider range of technologies and user preferences.

1 Accessible Repair Model

Universities, community colleges and other organizations are defining and setting web accessibility policies and practices. Current policies and practices are based on the minimal standards of Section 508 [1] with encouragement to implement the more extensive requirements of W3C Web Content Accessibility Guidelines [2]. Based on these practices developers usually first create web resources using the techniques and markup that they want, usually without any consideration for accessibility. Near the end of a development cycle, the developer uses an automated web accessibility checking tool to tell them of known or potential accessibility problems.

Most web accessibility checking tools are based on markup pattern matching to identify the markup used with the requirements of web accessibility requirements and include tools like Watchfire WebXact, Webaim Wave and other tools [3]. The reports can identify a few known accessibility problems like ALT text for images and LABELs on form controls, but most of the items in the report require the user to make extensive manual checks for accessibility that require too much time and are often beyond the capabilities of most web developers.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 230 – 233, 2006.

The current practice therefore usually improves the use of ALT text of images, but does little to improve the navigation structure or ability of users to restyle content for their own needs. For example if a developer does not use header (H1-H6) markup none of these tools will complain, even though headers are critical for providing a navigation structure. Current tools also are not designed to support administrators in understanding the state of functional accessibility of the web resources they manage, so verification of accessibility was difficult or impossible for them to assess.

2 Functional Accessibility

The goal of accessibility at UIUC is to make web resources more functionally accessible to people with disabilities. One of the key components to obtain this goal is the need for tools that would estimate compliance with the manual checks required in other accessibility analysis tools and in general provide authors with more information to help them evaluate the accessibility of their web resources. UIUC has developed a set of HTML Best Practices [4], the Functional Accessibility Evaluator (FAE) and the Web Accessibility Visualization Tool [5] to support developers in using accessible markup, estimating the use of best practices, and help developers visualize the accessibility of their resources. The tools use the following functional accessibility requirements defined in five major topics:

1. *Navigation and Orientation:* The use of html structural markup like headers (H1-H6), form control labels and other HTML structures for identifying the topics and relationships between topics in web resources. .
2. *Text Equivalents:* Images, video, audio and other multi-media object in web resources need text descriptions for people who cannot use the objects.
3. *Scripting:* Javascripting, embedded objects and other technologies that generate dynamic or interactive content.
4. *Styling:* The use of markup and styling techniques that support users restyling content for their own needs and content adapting to the capabilities of the browser rendering the resource.
5. *Standards:* The use of web standards insures that web resources can be rendered by the widest range of web technologies, which supports the basic principle of interoperability that is the heart of web technologies.

HTML Best Practices. The CITES/DRES HTML Best Practices [4] translate the requirements of Section 508 and W3C Web Content Accessibility Guidelines into markup requirements for implementing common web page features. Each best practice includes:

1. Explanation on the accessibility issues
2. HTML markup associated with the best practice
3. Relevant Section 508 and WCAG requirements
4. Examples of the best practices
5. Other resources

For example the best practices related to navigation include:

- Uniquely titling each web resource
- Indicating major/minor topics
- Accessible menus
- Accessible dynamic Menus
- Labeling and grouping form controls
- Indicate default and changes in languages
- Creating accessible links
- Using list markup
- Simple and complex data table markup
- Issues in using accesskeys
- Frames

Functional Accessibility Evaluator (FAE). One of the major problems after an organization defines a web accessibility policy and practices is estimating the implementation of the policy. The Functional Accessibility Evaluator (FAE) provides a means to estimate the functional accessibility of web resources by analyzing web pages and estimating their use of the best practices. The tool does not determine if a resource or a collection of resources is accessible or not, but provide statistics on the use of accessible markup.

Some examples of a test for accessibility:

1. *Unique Titles:* HTML Best Practices require the TITLE element and an H1 should be used to uniquely identify web resources. The FAE tool reports the percentage of pages in a web site that have unique TITLE elements and a matching H1 element on a page.
2. *Headers:* The use of headers is mostly ignored by other scripting tools. FAE provides information on the use of headers (H1-H6) by providing counts on the number of headers used on a page and percentage of content that is contained in headers. This allows web developers to see how much of there content is contained in headers and the distribution of header types in a web resource.

There are rules for testing each of the functional accessibility features of navigation, text descriptions, styling, scripting and the use of standards. The test results are linked to both the HTML Best Practices document and the Web Accessibility Visualization Tool for web developers to find out more information about the results. The HTML Best Practices provide information on the accessibility and markup issues with the functional accessibility requirement and the Web Accessibility Visualization Tool provides a means developers to graphically view the accessibility issues.

Web Accessibility Visualization Tool. The Web Accessibility Visualization Tool provides a means to create graphical views of functional web accessibility issues based on the HTML Best Practices. For example, with headers, the visualization tool shows a rendering of the web page with all non-header text grayed out so the header information is highlighted on the page. This allows a developer to see if all the major topic areas of a web resource have a header.

Another example is looking at overall document structure. If there are no headers, lists, or other structural content on a web page, then the visualization of the web page puts all the text of the web page into one paragraph. The developer is then asked how easy is it for them to use the web resource when all the content is in one paragraph.

The visualization tool helps to take developer beyond the current lists of accessibility checks of current web accessibility analysis tools and provides them with a means to visualize the accessibility problems.

Accessible Web Publishing Wizard for Microsoft® Office. Professional web developers can learn how to create accessible markup, but the most important web content in education is created by instructors. Most instructors have little understanding or interest in learning about the details of HTML, let alone the additional details associated with web accessibility. Therefore the tool that they use will have a tremendous impact on the accessibility of their web resources. The Accessible Web Publishing Wizard for Microsoft® Office provides them to use their most familiar authoring tool, Microsoft® Office, to create highly accessible web versions of the resources.

3 Conclusion

For functional accessibility to be achieved, web developers, instructors, students and all users need to embrace the benefits of accessible design principles. People need to value the additional options and control they have over the rendering and navigation of web content. As long as web developers view the needs of people with disabilities as something that does not benefit them or their users, the functional accessibility of web resources will remain marginal. Only when all users expect and appreciate the features of accessible design will there be a break through in making web resources more accessible to people with disabilities.

References

1. Section 508 Electronic Information Technology Accessibility Standards, http://www.access-board.gov/sec508/standards.htm
2. Chisholm, W., Vanderheiden, G., Jacobs, I. (2000) W3C Web Content Accessibility Guidelines, http://www.w3.org/TR/WCAG
3. iCITA Web Accessibility Evaluation and Repair Tools, http://cita.uiuc.edu/evaluation/
4. DRES/CITES HTML Web Accessibility Best Practices, http://cita.uiuc.edu/html-best-practices/
5. iCITA Software Tools for Accessibility, http://cita.uiuc.edu/software/
6. Illinois Accessible Web Publishing Wizard for Microsoft Office, http://www.accessiblewizards.uiuc.edu

A Proposed Architecture for Large Scale Web Accessibility Assessment

Mikael Holmesland Snaprud[1], Nils Ulltveit-Moe[1], Anand Balachandran Pillai[2], and Morten Goodwin Olsen[1]

[1] Faculty of Engineering, Agder University College, NO-4876 Grimstad, Norway
{Mikael.Snaprud, Nils.Ulltveit-Moe, Morten.G.Olsen}@hia.no
[2] SpikesourceIndia Pvt Ltd, Bangalore - 560034, India
abpillai@gmail.com

Abstract. This paper outlines the architecture of a system designed to demonstrate large scale web accessibility assessment developed in a European research project. The system consists of a set of integrated software components designed to automatically evaluate accessibility metrics for a large number of[1] websites and present results in a common report. The system architecture is designed to be maintainable, scalable, and extensible in order to facilitate further development of the tool. To meet these design criteria within a limited set of resources, an Open Source approach is adopted both for selecting, designing and developing the software.

Keywords: Software architecture, Web accessibility evaluation, free/open source software.

1 Introduction

Access for all to the Information Society is a declared key-goal for the European Union. With the wide-spread adoption of the Internet – in particular for access to government information and services, it is essential to secure access for all citizens. However, there are still digital barriers that make access to information difficult, especially for people with disabilities. Large scale web accessibility benchmarking can help to locate potential barriers and fuel the development towards Internet accessibility for all.

The EIAO (European Internet Accessibility Observatory) [1] project will establish the technical foundations of an Internet observatory for large-scale assessment of the accessibility of European websites.

The general specification of the demonstrator is defined in the Unified Web Evaluation Methodology, UWEM0.5 [2] which is based on WCAG1.0 [3]. To build a demonstrator implementation of UWEM0.5 we need to integrate different web assessment tools and aggregate the results into one coherent report. Experts often use

[1] The project is co-funded by the European Commission DG Information Society and Media, under the contract IST-004526.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 234 – 241, 2006.

several automatic tools to support their work of evaluating web accessibility. The results are then manually compiled into one report.

In this paper we outline some related work on web assessment components and architecture and elaborate on the chosen software components and interfaces in a proposed architecture. Finally, we discuss potential extended uses of the proposed architecture.

2 Related Work on Tools and Integrating Architecture

Most existing web evaluation tools are built to be self-contained and do not encourage integration with other tools. However, such automated tools often give a better coverage of accessibility aspects when used in combination. In [4] an architecture is proposed for integrating test tools to be used by practitioners who carry out web accessibility testing.

This problem has a range of common issues with the approach pursued in the EIAO project. Those issues include the difficulty to combine the reports from different tools both in terms of format and aggregating results, and the need to run tools individually and sequentially.

The proposed architecture in [4] outlines a common infrastructure that will allow tools to be installed, selected and customized. Those tools would share a common crawler and parser. Furthermore, seven tools are analysed in more depth for WCAG [3] coverage.

To classify the tools a taxonomy of capabilities is proposed. For results, a canonical format is referred to. EARL [5] will be addressed for this purpose as soon as EARL is fully defined by W3C.

In contrast with the architecture in [4], we base the tool interface on EARL and replace the indicated user interface with an import interface into the Observatory for further processing. We also need to make sure the integrated evaluation tools are at least partially conforming to UWEM [2].

Some early versions of our work on an architecture was first presented in [6] and later in more elaborated form dividing into components in [7]. An early presentation of measurements results is available in [8].

3 Outline of the EIAO Software Architecture

The architecture is based on a set of requirements derived from the original project proposal, a user requirement analysis and on the defined project Open Source policy as indicated in the next section.

3.1 EIAO and the Open Source Policy

The project management has decided to use the philosophy of OSS in the project. The policy is inspired by [9], by Bruce Perens. The following are the primary reasons for adopting this open approach:

A. Free access to source code, covering

- Availability: The software should be available free of charge so that others could use it without hindrances for research and improvement.
- Extensibility: The software created by the project should be open so as to allow extensibility by the use of plug-ins, adapters or other extension mechanisms.
- Transparency: The transparency of software used for measurements is essential to allow third party verification of the implementation. Also the software development process should be open and well documented to provide others a chance to verify the development methodology and tools.
- Participation: By defining a policy to use and develop OSS software, the project opens itself to participation from international developers and groups, who develop open/free software. This has helped the project to attract wider developer participation.

B. Free access to assessment results, covering

- Dissemination: Providing free access to the measurement data created by the project will help the project to reach out to a wider audience who can benefit from the results of the research.

- Verifiability: The results should be available to allow third party replication of measurements and verifying the metrics.

The above two aspects are crucial for validity of any measurement program.

C. Licensing. The project aims to provide all software developed by it under the GNU [10] General Public License (GPL). The publications of the project will be provided under the Creative Commons Attribution ShareAlike license. For a plug-in interface part of the software, the LGPL license or GPL with an exception is considered; which will allow the use of plug-ins released under proprietary licenses. The chosen licensing scheme has proved to simplify the collaboration within the project allowing the participants to focus on the research rather than on IPR discussions.

D. Open technologies. In addition to Python [11] for software development and Plone [12] for the project website the project uses a number of other OSS components. Those include PostgreSQL, Subversion, Apache, Trac, 3store, HarvestMan, RELAX-NG/Schematron/EARL and Openoffice.org.

EIAO reuses OSS components to reduce development lead time. The OSS development style allows the changes made to the components to be contributed back to the original projects.

The architecture is based on this policy and designed to meet the functional requirements.

3.2 Outline of the Architecture

The architecture is designed to facilitate the integration of existing and new accessibility evaluation tools into the observatory framework. The main components are shown below in Figure 1. A cron job will start the controller at regular intervals (monthly). The controller will then load the plugin configuration from the RDF repository, and start binding local and remote WAM components to the Scheduler/Dispatcher, based on the evaluation to be performed on the defined list of URLs in the URL directory.

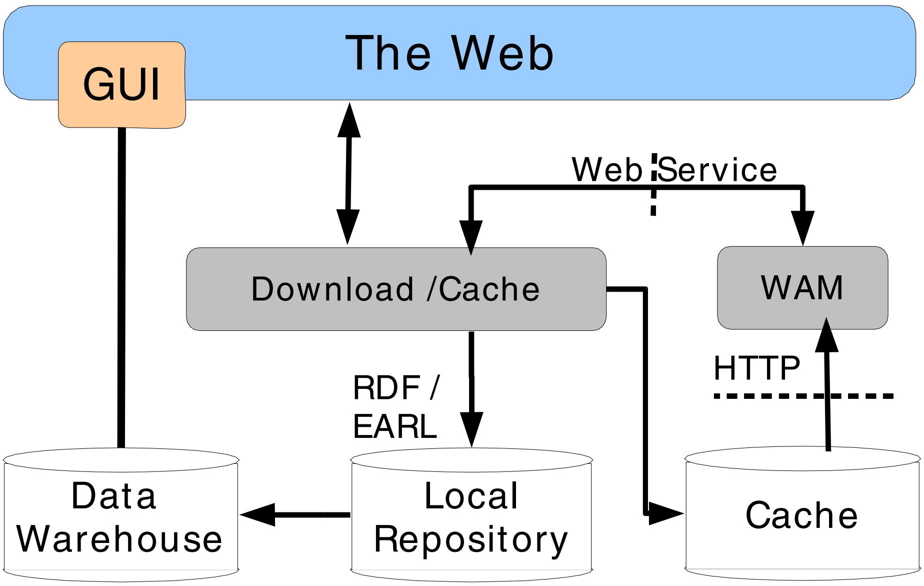


Fig. 1. The architecture of the Observatory planned for the first release

For the implementation of the architecture we will use the HarvestMan crawler [13], Relaxed [14] and Imergo [15] as basis for WAM plugins. In addition, we use 3Store on MySQL [16] as RDF repository, and PostgreSQL [17] as the basis for the Data Warehouse.

Information flow in the observatory. The information flow is structured as follows:

- The crawler downloads web page samples ta the http-cache repository.
- The downloaded web pages get their accessibility measured by internal/external WAMs.
- Subsequently, the sampling algorithm will choose the next sample to download. In release 1, this will be a uniform scheduling algorithm, but the scheduler will be candidate for a more advanced resource allocation algorithm in further releases.
- The results from the WAMs will be processed in a parallel queue-thread and later stored in the local RDF repository. At the same time the crawler will start downloading a new page / sample site.

Writing to the RDF repository and downloading from the Web are potential bottlenecks in the system. An advantage of having the information flow of the

Observatory as described above, is that the two bottlenecks of the system can run in parallel to improve the overall performance of the Observatory.

3.3 The Main Software Components Used in the First Implementation

Crawler. The key component in this architecture is the crawler, which is based on the Open Source web crawler HarvestMan [13]. HarvestMan was chosen because it is a good and mature web crawler that was easy to adapt to the system. Moreover, the HarvestMan developers were interested in actively supporting the EIAO project. HarvestMan has already been adapted to crawl according to EIAOs URL directory by the developer. HarvestMan has been further modified and integrated into the system to perform accessibility assessments on each downloaded web site and to store key use scenarios [2]. These scenarios are estimated by using the near uniform random sampling algorithm [18]. The web content is stored in a 3Store for further processing by web accessibility test modules.

Datawarehouse. A data warehouse is used in the architecture to support storage of large amounts of data and support flexible analysis from an on-line user interface. We also expect the hierarchical design of the databased schema in a data warehouse to facilitate analysis of the data.

PostgreSQL was chosen since it is the only Open Source DBMS with extensibility, table partitioning and bit-mapped indexes. Materialized view support is being built by the bizgres project. PostgreSQL performs very well for complex queries on large databases. The only other Open Source alternative is MySQL, which does not have these features.

Local repository. This component is responsible for storing the downloaded web pages, the URLs to crawl and also the results of the evaluation. The RDF repository will use the 3store RDF triplestore, and we will develop an EIAO RDF schema that can be used as an overall container for timestamped test runs containing EARL metadata, the URL directory, HTTP header information and other metadata extracted from the documents. Also data that does not fit nicely into EARL, like unique descriptions of the HTML deviations from Validate will be stored in the EIAO RDF schema. A formal representation of the EIAO RDF result repository also elaborated. The observatory in version 1 uses only one RDF repository, in order to make extraction of data as simple as possible for the data warehouse ETL tool.

The choice of a triplestore for the RDF repository was influenced by the co-operation with W3C on developing the EARL report language, since a large-scale EARL/RDF repository, able to collect data from different vendors, will be the first large-scale EARL consumer application. This may also be a driving force for vendors to use EARL.

Web Accessibility Metrics (WAMs). A Web Accessibility Metric (WAM) is a formal rule specifying how to make a statement about accessibility barriers of a given Web resource.

The description of a WAM consists of one or more procedures for how to measure characteristics of a Web resource and one or more procedures using the results of the measured characteristics to produce probability values for the violation of accessibility requirements derived from WCAG 1.0 / UWEM0.5.

In addition, the WAM will have one aggregation scheme (according to UWEM) that defines how to aggregate the probability values to produce a statement about an aspect of accessibility barriers of the Web resource. The current implementation gives two aggregated scores indicating the barrier probability and the change in barrier probability since the most recent measurement.

A set of Web Accessibility Metrics (WAMs) has been specified based on the tests described in UWEM0.5, checking for deviations from web standards (WCAG guidelines [3]) based on WCAG 1.0 checkpoints, with a view to conversion to WCAG 2.0 if possible. The first set of WAMs will be based on Relaxed/Schematron [14]. The first WAMs will address simpler forms of rules that can be expressed as Schematron rules. More complex rules may be implemented in some other language in a later release.

Cache. The Observatory must implement a web page repository to store versions of documents downloaded during the test runs and associated versioned tests and other metadata in raw form. It is important to store data unmodified to verify correct operation of the Observatory and as evidence for the data being tested.

Another important feature that should be considered is, facilities to avoid storing duplicates of the mirrored web documents. If the individual documents can be verified, by using digital signatures or hashing, we could save disk space by storing only a single version of each document downloaded. This is an important consideration, since storing 12 separate copies of test runs with all the (X)HTML data will exceed the current storage capacity of the Observatory.

To facilitate the checking of documents via WAM components or external tools, it is important that the web cache can be accessed via a transparent proxy, since most existing accessibility assessment tools will load missing web resources autonomously using the HTTP protocol. This will also apply to some of our own design bases, like the Relaxed validator.

4 Future Development of the Architecture

We have proposed a design suitable for integrating already existing software components to reduce the development and maintenance effort. We expect the architectures to evolve as the project matures, which should help to improve the performance. Interesting still unresolved issues include how to deal with the integration of potentially conflicting results from different tools or how to strike a balance between accuracy storage demands and sampling.

To make sure that the measurements are actually measuring barriers in a meaningful way, we have planned an improvement cycle involving systematic user testing. The test will be carried out with different disability groups on a set of web sites also assessed automatically. We plan to experiment with pattern recognition techniques and learning automata to identify the relations among user experience and automatically assessed web content. We also plan to improve the WAMs by extensive user feedback, to keep them updated with evolving Web technology. A more elaborate description of this approach is given in [19].

5 Conclusion

The proposed architecture builds on the results from a number of Open Source projects to establish a demonstrator for an Internet Accessibility Observatory.

Much effort was initially spent on defining the requirements and on code analysis to select the most suitable code bases. For the first implementation of the architecture we use the HarvestMan crawler [13], Relaxed [14] and Imergo [15] as basis for WAM plugins. In addition, we use 3Store on MySQL [16] as RDF repository, and PostgreSQL [17] as the basis for the Data Warehouse.

The results from the prototype implementation indicate that adapting and integrating existing tools rather than building a system from scratch has been very suitable for the task. This choice has also, implemented by the highly talented development team, resulted in an extensible architecture for large scale web accessibility evaluation.

We think that a similar architectural approach can be useful also in related research projects and plan to explore this potential further in future research.

References

1. European Internet Accessibility Observatory - http://www.eiao.net
2. Unified Web Evaluation Methodology version 0.5 - http://www.wabcluster.org/uwem05/
3. Web Content Accessibility Guidelines version 1.0 - http://www.w3.org/TR/WAI-WEBCONTENT/
4. Englefield, P., Paddison, C., Tibbits, M., Damani, I.: A Proposed Architecture for Integrating Accessibility Test Tools, IBM Systems Journal, Special Issue on accessibility, Vol 44, No. 3, (2005) 537-555, - http://www.research.ibm.com/journal/sj/443/englefield.html
5. Evaluation and Report Language - http://www.w3.org/TR/EARL10-Schema/
6. Snaprud, M.H., Jensen, C.S., Ulltveit-Moe, N., Nytun, J.P., Rafoshei-Klev M. E., European Internet Accessibility Observatory (EIAO), eGEP - Getting full Impact of eGovernment - launching a European Measurement Framework-, Vienna, Austria, (2006)
7. Ulltveit-Moe, N., Gjøsæter, T., Snaprud M. H.: An Open Architecture for Large Scale Monitoring of Web Accessibility - Workshop on Web Accessibility and Metamodelling, Grimstad Norway, (2005)
8. Snaprud, M.H., Ulltveit-Moe, N., Granmo, O.C., Menk, M.E., Wiklund, A., Sawicka, A.: Quantitative Assessment of Public Web Sites Accessibility - http://www.etsi.org/cce/proceedings/7_3.htm
9. Perens. B.: Global Technology Policy Institute - http://techp.org/
10. Stallman, R.: GNU General Public License (GPL) - http://www.gnu.org/copyleft/gpl.html
11. van Rossum, G.: Python - http://python.org/

12. Plone: A user-friendly and powerful open source Content Management System - http://plone.org/
13. Pillai, A.B.: HarvestMan (2005) - http://harvestman.freezope.org/
14. Nalevka, P.: Relaxed HTML Validator2005 - http://badame.vse.cz/validator/
15. Velasco, C.: Imergo 2005 - http://www.fit.fraunhofer.de/projekte/imergo/index_en.xml
16. Riddoch, A., et al: 3store2005 - http://sourceforge.net/projects/threestore/
17. PostgreSQL Global Development GroupPostgreSQL2005 - http://www.postgresql.org
18. Henzinger, M.R., Heydon, A., Mitzenmacher, M., Najork, M.: On Near-Uniform URL Sampling (2005), http://research.compaq.com/SRC/mercator/papers/www9/
19. Craven, J., Snaprud, M.H.: Involving Users in the Development of a Web Accessibility Tool, http://www.ariadne.ac.uk/issue44/craven/intro.html

People with Disabilities: Quality of Web Accessibility
Introduction to the Special Thematic Session

Dominique Burger

Association BrailleNet
INSERM U592
9, quai Saint Bernard, 75252 Paris cedex 05, France
`dominique.burger@snv.jussieu.fr`

Abstract. This session has been design as to discuss several and various aspects involved in the design of quality and to examine how e-accessibility may impact this general process.

1 Introduction

Making the ever-growing multitude of Web sites (both public and private) accessible within the next few years will be a major endeavour across Europe and throughout the world. This implies a considerable effort in R&D and Industry for providing a variety of technical e-accessibility solutions involving access and assitive technology tools as well as methodologies for the creation and implementation of accessible content and evaluation, all in concert with legislation and standards.

There is now a matter of evidence that e-accessibility shall be taken into account early in the design and implementation process starting with technical specification, and the choice of the appropriate development tools. Moreover e-accessibility should be incorporated into standards, referenced in legislations, and systematically addressed in public procurement. The professional developers should have been taught the Design for All principles and have had a practical experience of them. This would affect the final approval of goods and services by purchasers, on the basis of conformity certification.

In other terms, e-accessibility should simply be considered as an element in Web Quality Management, and become a part of the decisional and technological process.

This would certainly represent a cultural change, but would dramatically vitalise the industrial process which would have to produce better customisable solutions and fully interoperable products.

The ambition of this session is not to cover all the aspects involved in this idea, but at least to foster an in-depth reflection on some of them.

The session is organized on behalf of the European Specific Support Action Support-EAM[1].

---

[1] Support-EAM is an IST funded project (FP6-2003-IST-2-Support Action 004754): Support the Creation of an eAccessibility Mark (`www.support-eam.org`).

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 242 – 244, 2006.

2 Overview of the Session

User-centered design is widely recognised as a key for success in the development of human-computer interfaces. Nevertheless taking into account the needs and preferences of people with disabilities raises specific questions that standard methodologies and guidelines may ignore, making necessary to widen preliminary studies with users, involving users with disabilities at an early stage of the design. This question is discussed by M.L. Guenaga, J. Oliver and A. Barbier in their paper "Accessible Interfaces to Empower the User" in which they report on two studies conducted with blind uses for better understanding some possible interaction barriers and proposing a path towards necessary improvements. The study was conducted on the basis of a Web questionnaire an prototype testing, bringing concrete suggestions.

Accessibility guidelines provide general and useful directions to the designers. Nevertheless, M. Schrepp and P. Fischer underline that these guidelines may lack the necessary guidance and details for practical design. For instance, the W3C/WAI WCAG1.0 say that " product functions shall be executable from a keyboard". The authors underline that it is not sufficient and that, in addition, a user must be able to navigate a page using the keyboard with acceptable performance. In their paper "A GOMS Model for Keyboard Navigation in Web Pages and Web Applications" they discuss how these models can be used to compare different ways to operate a web page or web application and improve the efficiency of users with disabilities.

Certification has a potential for fostering e-accessibility. In the paper "Towards Web Accessibility Certification: The findings of the Support- EAM project", D. Burger presents the divergent opinions of different groups of stakeholders concerning Web accessibility conformity certification. These opinons were expressed formally in a CEN Workshop Agreemeent (CWA) identifying three possible schemes for Web accessibility certification, in Europe. These schemes are 1) inspection according to the ISO/IEC 17020, 2) certification according to the EN 45011, and 3) Supplier's Declaration of Conformity according to the ISO/IEC 17050. Nevertheless, most stakeholders recognise the need for global standards and a harmonised implementation of the W3C/WAI Web Content Accessibility Guidelines. This might be the mission of an Institute that would produce a normative document on Web accessibility conformity assessment and create a Web accessibility Quality Mark.

In the paper "Web Accessibility Conformity Assessment - Implementation Alternatives for a Quality Mark in Austria", M.L. Leitner, K. Miesenberger, D. Ortner and C Strauss examine how the recommendations made by the CEN Workshop on Web Accessibility certification can be practically implemented at national level, taking the example of Austria. The authors draw four basic senarios based on the CEN Workshop Agreement. They evaluate them against six criteria, namely complexity, cost, dependence, flexibility, impartiality, and time. They also consider the legal framework in Austria, in order to propose a method for analysing the advantages and drawbacks of various implementation alternatives and to make the appropriate decsisions for their implementation.

F. Aslaksen, F. Fardal and M. Snaprud, analyse the "The role of benchmarking in concerted actions to increase accessibility". A experiment was made in Normay in 2001 for measuring the conformity of websites to requirments and making the results public, and thereby "reward" websites of good quality. This was reconducetd year after year and a quality mark was created for supporting such policy. The authors demonstrate the qualitative impact of such a policy on e-accessibility. They propose to extend it at European level by means of a European Internet Accessibility Observatory. This is the objective of an European funded project (EIAO) whose aim is to develop a prototype technical machinery for large-scale Web accessibility conformity assesment and publication.

The European Community and several Member states have recognised that Design for All should be part of any design curriculum, as a key factor in building a better inclusive society. This requirement is more and more considered in architecture curricula for the building of spaces that meet the needs of all people, young and old, abled and disabled. This should also affect engineering courses. K. Miesenberger and D. Ortner report the successful experience of B.F.W.D. - Accessible Web Design (Barrierefreies Web Design) in Austria as a way of "Raising the Expertise of Web Designers through Training". This course was developped by Institute Integriert Studieren at the University of Linz, designed for post graduate students and practitioners working in the field of web design. It is worth noting that the course is e-learning based and designed to be accessible to students and teachers with disabilities, as well.

3 Conclusion

This sessions shows that Web accessibility could irrigate technological, industrial and educational policies for producing innovating and high quality solutions. If this ideal were realised the additional cost of Web accessibility – and e-accesibility in general – would be absorbed in the overall cost of global quality. But the added-value of accessibility would certainly be beneficial to all.

Towards Web Accessibility Certification: The Findings of the Support-EAM Project

Dominique Burger and Pierre Guillou

INSERM U592 – Université Pierre et Marie Curie
Association BrailleNet
9, quai Saint Bernard, 75252 Paris cedex 05, France
{Dominique.Burger, Pierre.Guillou}@snv.jussieu.fr

Abstract. This paper presents the results of a European Specific Support Action whose objective was to explore the possibility to certify the accessibility of Internet services and the possible implications for a eAccessibility Quality Mark. The project initiated and conducted a CEN workshop which produced a CEN Workshop Agreemeent (CWA) identifying three possible schemes for Web accessibility certification, namely inspection according to the ISO/IEC 17020, certification according to the EN 45011 and Supplier's Declaration of Conformity according to the ISO/IEC 17050. The CEN workshop also recommended the creation of an Institute whose objective would be to harmonise the interpretation and implementation of the W3C/WAI Web Content Accessibility Guidelines (WCAG) in Europe. The Institute would create, promote and maintain a Quality Mark.

Keywords: eAccessibility, Web accessibility, Conformity, Certification, Supplier's declaration, Quality Mark.

1 Introduction

Since 2001, e-Inclusion has become a key objective of the overall European Social Inclusion Strategy. The European Community has established objectives towards e-accessibility within the i2010 – "A European Information Society for growth and employment" [1]. Concerning Web Accessibility particularly, guidelines have been formulated by the W3C Consortium which are recognised worldwide, the Web Content Accessibility Guidelines (WCAG) [2]. In reality, several studies have evidenced that the WCAG are quite poorly implemented, like the one published by the UK Cabinet Office in November 2005 [3].

Certification is one of the instruments that are considered as having a potential for improving the current situation, and supporting the legislations that mandates accessibility [4].

In several countries this assumption is supported by the experience of non governmental organisations which have created conformity labels based on the WCAG guidelines. Among other, one can mention AccessiWeb, Blindsurfer, Drempelvrij, See it Right-Useability, Sello de Accessibilidad [5,6,7,8,9]. These labels are based on private agreements between service suppliers wishing to demonstrate the

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 245 – 252, 2006.

quality of their products and reinforce their public image, and organisations being accountable of technical verification for which the suppliers have not the expertise. These initiatives have improved the clarity and the transparency in the attribution of accessibility labels and have contributed to make accessibility recognised as a necessary component in the development process.

Nevertheless, the success of these labels raises the question of a real control of their quality and their harmonisation. If this question were not solved, any policy for a large-scale implementation of the W3C WCAG would be difficult, due to their fragmented interpretation. In this context, the Council of Europe stated in 2003, that a Quality Mark for eAccessibility should be set up, and the European Commission recommended certification as an instrument for fostering eAccessibility, in the Communication released on 13 September 2005 [10,4]. A mandate to European standardisation bodies concerning eAccessibility was released in December 2005 by the European Commission [11].

A specific Support Action was set up in order to explore the possibility to certify the accessibility of Internet services in Europe. This action is called Support-EAM[12]. It was conducted by organisations from 7 countries[1], and coordinated by the association BrailleNet, France, from October 2004 to July 2006. The two main objectives of Support-EAM were:

- To develop a harmonised methodology used for assessing Web Accessibility (single-site or large-scale).
- To recommend requirements for a Web Accessibility certification scheme.

2 What Is Certification?

Service certification is a standard procedure for garanteeing that a service is in conformance with requirements specified in a norm or a normative document, on an objective basis. Service certification is a voluntary action undertaken by providers that outsource the necessary operations to a third-party organisations, called certification body, in order to bring garantees to customers and users. The certifying organisation uses inspections, that they carry out themselves of with help of external organisations. The guaranty covers a certain duration. This guaranty is often associated with a Certification Mark in order to make it more visible.

Service certification and the related procedures are defined in standards such as ISO/IEC Guide 28:2004, ISO/IEC Guide 65, ISO/IEC 17020:1998. The certification body shall be competent, respect impartiality and transparency. It is accountable for what it certifies. It shall also insure that normative documents are valid and accepted by stakeholders. If needed it shall organise their update with the parties concerned.

In many domains, service certification has proven to be an efficient instrument helping companies to build up confidence with their customers and improving their competitiveness.

---

[1] The partners are Association BrailleNet (France), Fundosa Teleservicios (Spain), Bartimeus Accessibility Foundation (Netherlands), Dublin City University (Ireland), Universität Linz (Austria), Katholieke Universiteit Leuven Research & Development (Belgium), AccessInMind Ltd (United Kingdom).

3 Supplier's Declaration of Conformity

In order to alleviate control procedures and reduce the costs another conformance certification system has been set up by companies, based on Supplier's Declaration of Conformity (SdoC). Suppliers may decide to adopt this scheme if they are convinced that their image and expertise are good enough to allow them to save the cost of third party-certification. They remain accountable for their declarations. SdoC shall be complemented by market surveillance and respect the ISO/CEI 17050. This approach has been successfully applied in several sectors like telecoms or electrical goods.

4 A Demand for Web Accessibility Certification

The Support EAM has conducted two on-line digital surveys where a wide spectrum of stakeholders expressed a need for a certification scheme [12].

4.1 Stakeholders' Needs for a Quality Mark

A first survey was conducted in February-March, 2005, whose objective was to collect opinions about a European Quality Mark. About 450 answers from more than 20 countries were received. They represent a great variety of stakeholders with a strong representation of public organisations and Web designers. The results definitely highlight the needs for clearer communication on Web accessibility and the set up of a methodology and a verification process to put into concrete the WCAG:

- 80% of respondents pay attention to Web accessibility of their sites but only 35% have checked it: most people know of Web accessibility but have no clear ideas of what to do.
- 75% of respondents want methodology and criteria.
- 80% of respondents think that to get a Quality Mark for their Web site is necessary before acceptance from their site builder: the Quality Mark is seen as a security and quality issue.
- Stakeholders also consider that a Quality Mark for Web Accessibility would improve the image of the owner of the website (81%) and trust the accessibility of the Web site (84%). They think that an accessible Web site reaches a larger audience and therefore increase the Return On Investment (70%).
- 72% of the respondents have a budget for accessibility. Furthermore, 34% of them are willing to pay for an ongoing evaluation of their website. Among them, around 50% were even proposed an estimated annual budget for that whose average is 2200 €/year.

4.2 Stakeholders' Needs for a European Authority

A second on-line digital survey was conducted in February-March, 2006, whose objective was to collect opinions about a European Authority for certification. The 241 respondents who answered the questionnaire originated from more than 13 different European countries.

A striking result of the survey is that third party certification is regarded as a more trustful certification method than Supplier Declaration of Conformance. A majority of respondents is in favor of 3rd party certification, half of them accepting to pay for this. For most respondents avoiding fragmentation of interpretation of web accessibility guidelines and harmonisation should be the key activitiy the authority (Table 1).

Table 1. Answers to the question: In your opinion, what is the most important activity of the European Authority concerning the operation of the certification scheme?

	Total	Perc.
Avoiding the fragmentation of the interpretation of European and W3C and other international guidelines for web accessibility	164	68%
The harmonization of quality marks	134	56%
Drafting a normative document setting criteria to web accessibility	129	54%
Registration of website which have been attributed the quality mark	121	50%
Development of a European quality mark	109	45%
Contracting parties issuing this quality mark	49	20%
Effectuating market surveillance	33	14%
Total	739	

The international role of the authority is considered most important, while testing should be made by a local authority.

The vast majority of respondents share the opinion that as a European Authority the best way to support an adequate quality mark is to encourage the use of objective criteria and good practices concerning web accessibility.

These results confirm those of the on-line consultation from the Europa Commission open to all interested organisations and individuals to participate, in January-February 2005.

5 Can On-Line Services Be Certified?

Several answers can be made to the objection that Web sites are changing permanently making any evaluation obsolete immediately after it is completed:

Firstly, this problem is not specific to on-line services, but rather common for many services, like services welcoming the public, elevators, touristic sites, for instance. Nevertheless, it is accepted that certification can guarantee their quality;

Secondly, on-line services may be quite stable, like some administrative services, for instance;

Finally, there are solutions to the problem of changeability of Internet services: the frequence of controls shall be adjusted according to the stability of the code; controls can be preformed remotely and randomly, using automatic surveillance tools; the services are continuously exposed to public so that complaints can be collected in real-time. Such features make possible on-going surveillance and corrections. Several experiences of accessibility labeling scheme have taken profit of them, imposing to Web site suppliers to address accessibility issues in a limited period of time.

6 Requirements for Normative Document

Normative documents shall provide precise criterias and methods making possible to verify them on an objective basis. They allow for the catacterisation of the services to be controlled. They can refer to a technical norm, or to widely accepted recommendations. In addition to technical specifications, normative documents specify how to inform customers and users about the conformity assesments and the guaranties that are given, and - possibly - commitments concerning the service and reparations made in case of non-conformity.

Concerning Web accessibility, no international norm exists. But the WCAG are officially reognised by European countries. Thus, a normative document for Web accessibility certification should be based on these recommendations.

National organisations have undertaken the adaptation of these recommendations into national reference documents, norms, or technical rules for implementing the laws. One can mention, for instance:

- Spain: Norm developed by AENOR : "Requisitos de Accesibilidad para Contenidos en la Web", UNE 139803:2004;
- France: AccessiWeb, developed by BrailleNet, submitted to public comments in 2003 and incorportated in the interoperability reference documents of the administration in 2004;
- Italy: Decree of 8 July 2005 “Requisiti tecnici e i diversi livelli per l'accessibilità agli strumenti informatici”, (technical rules) published in the Official Journal n. 183 8 August 2005;
- Pays-Bas: Normative Document Drempelvrij prepared by a normative commission;
- Royaume-Uni: “PAS 78: Guide to Good Practice in commissioning Accessible Websites”, developed by BSI British Standards.

Nevertheless, nor the WCAG, nor the national adaptations, can constitute by themselves the technical reference for certififying a Web site, because they do not contain methods for deciding the conformity on a uniquely interpretable, repeatable basis. This is why the various labeling schemes had to create their own documents in order to make their assessments as objective as possible.

7 The Web Accessibility Certification CEN Workshop

Support-EAM organised the concertation with stakeholders within a CEN/ISSS Workshop (WS/WAC) in order to establish a first level consensus for a certification scheme and a Quality Mark [13]. The Workshop has elaborated a scheme which was adopted as a CEN Workshop Agreement (CWA) after a 60 days public comment, on April 6th, 2006, according the the CEN/ISSS policy.

The CWA identifies 3 schemes reflecting the preferences that were expressed:

- Inspection providing audit reports, according to ISO/IEC 17020;
- Service certification by accredited organisations, according to EN 45011;
- Supplier’s declaration of conformity, according to ISO/IEC 17050.

The CWA proposes a reasonable way for existing labeling schemes to improve and migrate towards standards, according the model fitting the best their needs.

An Institute would act as a central authority whose mission would be to organize, harmonize and control practices in liaison with the international bodies concerned. The Institute would be the owner of a Quality Mark (or Conformity Mark) and would develop services for promoting Web accessibility.

8 A European Authority for Web Accessibility Accessibility Assessment

The organisations which operate web accessibility quality marks for web accessibility in Europe would join their efforts for the creation of the institute whose role would be:

- avoiding the fragmentation of the interpretation of guidelines for web accessibility;
- the harmonisation of web accessibility quality marks;
- drafting a normative document, in line with the adopted version of the WCAG;
- development of the web accessibility quality mark;
- establishing agreements with parties issuing the web accessibility quality mark;
- registration of websites which have been attributed the web accessibility quality mark;
- effectuating market surveillance;
- creating an active dialogue with user groups.

The Authority should guarantee its independence, and should be submitted to quality controls.

9 Roadmap Towards a European Certification Scheme

The road towards a European certification scheme should go through the following steps:

- Creation of the institute by organisations runing Web accessibility labels;
- Preparation of a normative document;
- Creation of a web accessibility quality mark and adoption of a convergence plan of the different labels towards this mark;
- Establishing relationships between the institute and the standardisation bodies;
- Promoting services in relation with the web accessibility quality mark, respecting independence principles;
- Establishing agreements with organisation willing to use the mark.

10 Conclusion

User and consumers organizations, industry and governments agree that the accessibility of technology is a significant and relevant issue that shall be addressed seriously and coherently.

Accordingly, in its Communication on eAccessibility, September 13th, 2005, the European Commission proposes a set of policy actions that should foster e-accessibility. It calls on Member States and stakeholders to support voluntary positive actions to make accessible ICT products and services far more widely available in Europe. It mentions certification as an instrument [4].

All see the accessibility of technology as a large-scale problem concerning many business domains and hitting a growing potential market. They consider that regulation, standards, e-procurement and some forms of recognized certification should help towards a society more beneficial to all citizens, matching the ethical objectives of democracies.

Also the work done by the W3C under the Web Accessibility Initiative is recognized as a basis for building up Web Accessibility, particularly the WCAG.

End of 2005 an EC mandate has been given to the European Standardisation Organisations to come with a solution for common requirements and conformance assessment.

This context is encouraging the creation of a certification system for Web Accessibility designed to work at European level. The experiences that organisations have conducted in several countries over the past few years constitute a favourable basis for that.

The creation of an institute would offer an opportunity to bring together competences and energies, and stop the fragmentation into multiple schemes, whose paradox would be to make any further progress towards accessibility difficult in Europe.

The work undertaken by Support-EAM and the CEN Workshop clearly show that such an authority could integrate different existing schemes supporting voluntary certification and serving the interest of all the parties concerned. This authority would make collective expertise visible and provide a space where to develop it and improve it collectively. Organisations having a solid experience in labelling Web sites should constitute the core of this authority that should also take on board all the interested stakeholders, in order to capitalise the lessons from different experiences, which have produced a real expertise that should be promoted at European level.

Acknowledgments. Support-EAM is an IST funded project (FP6-2003-IST-2-Support Action 004754).

References

1. i2010 - *A European Information Society for growth and employment*: `http://europa.eu.int/information_society/eeurope/i2010/index_en.htm`
2. W3C/WAI: `www.w3.org/WAI`
3. UK Cabinet Office, *eAccessibility of public sector services in the European Union: http://www.cabinetoffice.gov.uk/newsroom/news_releases/2005/051124_eaccessibility.asp*
4. *Communication from the Commission to the Council, the European Parliament, the Economic and Social Committee, and the Committee of Regions* (13 September 2005): `http://europa.eu.int/information_society/policy/accessibility/regulation/com-ea-2005/a_documents/cec_com_eacc_2005.htm`
5. Accessiweb: `www.accessiweb.org`

6. BlindSurfer: http://www.blindsurfer.be/
7. Drempelvrij: www.drempelvrij.nl/waarmerk
8. See it Right-Useability: http://www.abilitynet.org.uk/content/oneoffs/seeitright.htm
9. Sello de Accesibilidad: www.teleservicios.com/SRV/srv_auditoria.asp
10. Council Resolution on *eAccessibility - improving access people disabilities to Knowledge Based Society* (doc. 5165 03): http://europa.eu.int/comm/employment_social/knowledge_society/res_eacc_en.pdf
11. EC mandate given to the European Standardisation Organisations: http://europa.eu.int/information_society/policy/accessibility/regulation/pubproc/index_en.htm
12. Support EAM: www.support-eam.org
13. Comité Européen de Normalisation (CEN): www.cenorm.be

Raising the Expertise of Web Designers Through Training – The Experience of BFWD – Accessible Web Design (Barrierefreies Webdesign) in Austria

Klaus Miesenberger and Daniela Ortner

University of Linz, Institute Integriert Studieren,
Altenberger Str. 69, 4040 Linz, Austria
{klaus.miesenberger, daniela.ortner}@jku.at

Abstract. Over the last years a well elaborated body of knowledge in "Web Accessibility" has become available. Awareness and in accordance legal directives today ask for application of this knowledge. The BFWD post graduate course, a comprehensive university course on accessible web design, is a pro active reaction to this an increasing demand.

1 Introduction

The Institute Integriert Studieren at the University of Linz developed the course "Barrierefreies Webdesign" [1] what was partly funded by the European Social Fund (ESF) [2] and the Federal Ministry for Education, Science and Culture (bm:bwk) [3]. This course is designed for post graduate students and practitioners working in the field of web design. The course is e-learning based and designed in an accessible way for students and teachers.

After successfully implementing the curriculum, getting approved the course by the university board and preparing the first parts of the content the course started in October 2005 with 19 students coming from Austria and Germany. Most of the students work in web design companies which are confronted with the need to develop accessible web pages for their clients. by enclosing a check (mail orders only).

2 Legislation

Web Accessibility in Austria is regulated by two laws:

– The e-Government-Law [4] that has been enacted in 2004 states that websites offered by public authorities have to follow international Web Accessibility standards in order to allow access without barriers for people with disabilities. Following this law, web pages of public authorities in Austria have to be adapted by the beginning of the year 2008.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 253–257, 2006.

– Furthermore the so called “Behindertengleichstellungsgesetz” (anti discrimination legislation) was put in force at the beginning of 2006. This law guarantees the right of access to all areas of the public sector including web pages [5].

Due to this new legal framework an increasing demand for courses can be expected.

3 Development of the Curriculum

The development of the curriculum has formally been assigned to the committee responsible for computer science studies at the University of Linz. Based on the drafts delivered by the Institute Integriert Studieren and according to the intensive communication and discussions with the committee, the curriculum has been improved through several iterations. The curriculum has a total sum of 44 semester hours and is divided into the following six modules:

– Technical Fundamentals
– Assistive Technologies
– Guidelines and Laws
– Accessibility
– Design and Usability
– Practical Experience

Each of the modules contains appropriate lectures. Mandatory attendance hours are reduced to a minimum during the whole course as the main parts of the content are taught using an online e-Learning system. This was intended and also accepted as and invitations for experts on the job.

4 Inclusive Education

The goal of the postgraduate course was not only to teach people about Web Accessibility, but also to design the course itself in a way that people with disabilities are able to attend as well as to teach in the course. According criteria had to be taken into account when selecting and adapting the system. Several e-Learning platforms were analysed including commercial as well as non-commercial systems. It became obvious that commercial systems are at the one hand not affordable for the project, and on the other hand do not comply to the requirements regarding Web Accessibility. As the source is not available they cannot be adapted. The decision was between three non-commercial systems: ATutor [6], Ilias [7] and Moodle [8]. Moodle was selected for several reasons: it is open source, it fulfils most of the accessibility criteria, it seems to be adaptable without too much effort, and it allows content developers to use the tools they are familiar with (as for example MS Word for lecture notes). The following changes and features have been implemented to increase the accessibility and usability:

– Using relative instead of absolute sizes
– Inserting skip-links
– Inserting alt-attributes and appropriate values
– In forms, associating labels with their controls
– Developing an accessible chat

The adapted system now fulfils most priority 1 and 2 criteria of the Web Content Accessibility Guidelines 1.0 (WCAG 1.0) [9] of the Web Accessibility Initiative (WAI) [10]. To make self-studying easier, a lecture-like presentation tool for content including slides and synchronised speech with the additional possibility to navigate through slides and chapters has been developed. The first attempt was using the Synchronized Multimedia Integration Language (SMIL) [11], more precisely the Timed Interactive Multimedia Extensions for HTML (HTML+TIME) [12], which are based on the XHTML+SMIL language profile [13] of the Synchronized Multimedia Integration Language and add timing and media synchronisation support to HTML pages. But there were some problems that came along with using HTML+TIME: the use of streaming media is not supported, and the only supported browser is Microsoft Internet Explorer. In a second attempt, a prototype using JavaScript has been developed and is in use now. Using JavaScript does not make the application fully accessible according to the Web Content Accessibility Guidelines, but it is assumed that people with disabilities joining the course use modern assistive technologies. Our tests involving experienced blind computer users have shown very positive results.

At the beginning of the postgraduate course, a face to face meeting of 2 days was organised. This was used for an introduction to the course, for social contacts and a good start of co-operation and for an introduction to the e-learning system. Three face to face meetings are organised each semester. The last one includes exams.

Today blind, visually handicapped and deaf students attend the course. Also three blind and one severely mobility impaired teacher successfully use the system for delivering their courses. This shows that an adequate solution has been found.

5 Future Plans

After successfully starting the course in Austria first considerations and negotiations have been started to offer the course also in other countries and languages. This of course asks for localizing the content. Such co-operations should guarantee the offer in the long run and an efficient and cost effective organisation and updating.

The curriculum is designed following a modular approach, which allows that single lectures can be changed and also used also for other teaching purposes. Some lectures could for example be easily integrated into the regular lectures for computer science students or to provide training for interested companies and organisations.

As the e-Learning system Moodle is developed by a big community, there is a constant progress in functionalities. To keep the system up to date, it will be necessary to port the adaptations that were made regarding to Web Accessibility also to newer releases of Moodle. Discussions with the community about integrating our adaptations into their development and therefore also into new releases would not only be very useful for us, but also for all other organisations providing e-Learning courses via Moodle.

As soon as version 2.0 of the Web Content Accessibility Guidelines (WCAG 2.0) [14] becomes an official recommendation of WAI, the current installation of the e-Learning system has to be adapted according to the new specifications.

Additionally, also the contents of the course that deal with the guidelines have to be adapted in order to follow the new recommendation.

The Institute Integriert Studieren is currently preparing another postgraduate course which will use the same system for teaching: "Assistec - Assistive Technologies" [15] focuses on consulting and the development of the acquisition-process of Assistive Technologies in practice. It represents an interdisciplinary innovative program including technical, social, economical and medical aspects. The curriculum of this course became accepted by the university at the end of 2005. The first course will start in October 2006.

Acknowledgement

This project has been supported by the Austrian Federal Ministry for Education, Science and Culture (bm:bwk) and by the European Social Fund (ESF).

References

1. "Barrierefreies Webdesign" project homepage, online available on May 13, 2005 at http://www.bfwd.at
2. European Social Fund (ESF) online available on May 13, 2005 at http://www.esf.at
3. Federal Ministry for Education, Science and Culture (bm:bwk) online available on May 13, 2005 at http://www.bmbwk.gv.at
4. E-Government-Law (E-GovG), 2004, online available on May 13, 2005 at http://www.parlament.gv.at/pls/portal/docs/page/PG/DE/XXII/BNR/BNR_00149/fname_014980.pdf
5. Behindertengleichstellungsgesetz - Regierungsvorlage (government bill), 2005, online available on May 13, 2005 at http://www.parlament.gv.at/portal/page?_pageid=908,848700&_dad=portal& _schema=PORTAL
6. ATutor Homepage online available on May 13, 2005 at http://www.atutor.ca
7. Ilias Homepage online available on May 13, 2005 at http://www.ilias.uni-koeln.de/
8. Moodle Homepage, online available on May 13, 2005 at http://www.moodle.org
9. Web Content Accessibiliy Guidelines 1.0, W3C Recommendation, 1999, online available on May 13, 2005 at http://www.w3.org/TR/WCAG10
10. Web Accessibility Initiative (WAI) online available on May 13, 2005 at http://www.w3.org/WAI

11. Synchronized Multimedia Integration Language (SMIL) online available on May 13, 2005 at http://www.w3.org/AudioVideo
12. HTML+TIME online available on May 13, 2005 at http://msdn.microsoft.com/library/default.asp?url=/workshop/author/beh aviors/reference/time2/htime_node_entry.asp
13. XHTML+SMIL, W3C Note, 2002 online available on May 13, 2005 at http://www.w3.org/TR/XHTMLplusSMIL/
14. Web Content Accessibility Guidelines 2.0, W3C Working Draft, 2004 online available on May 13, 2005 at http://www.w3.org/TR/WCAG20
15. Assistive Technologies project homepage online available on January 23, 2006 at http://www.assistec.at

The Role of Benchmarking in Concerted Actions to Increase Accessibility

Finn Aslaksen[1], Frank Fardal[2], and Mikael Snaprud[3]

[1] Vista Utredning AS, C.J. Hambros plass 5, 0164 Oslo, Norway
finn.aslaksen@vistautredning.no
[2] Norge.no, Njøsavegen 2, 6863 Leikanger, Norway
Frank.Fardal@norge.no
[3] Faculty of Engineering, Agder University College, Grooseveien 36 NO-4876, Norway
mikael.snaprud@hia.no

Abstract. This paper covers "soft measures" that can be used in a "stick and carrot" strategy to improve accessibility to web sites. Two examples of benchmarking are presented. The first is the benchmarking of 700 Norwegian public web sites that has been carried out three times the last years, giving experiences on how to carry out the analyses and how to use the results for awareness raising and stimulation to improvement. The second is the planned European Internet Accessibility Observatory, EIAO[1]. This will be capable of large-scale web assessment and may speed up the effects of benchmarking as a tool for awareness raising.

1 Introduction

Universal design and access for all has recently received much attention both in public administration and in the private sector. In all fields lawmaking and the development of standards and regulations play an important role. Several countries have introduced anti-discrimination laws protecting disabled from discrimination, and there is a continuous work going on to improve standards and regulations (like the building regulation).

This is also the situation for the accessibility to the Internet. However, many countries have so far not agreed upon an anti discrimination law, and standards are not compulsory. Luckily there are also some other measures available to speed up accessibility. These may be regarded as supportive actions to the use of law and compulsory standards and should be included in the sticks-and-carrot-toolbox of government supported action to increase accessibility to the Internet. Such carrots may also be used before the sticks (law and compulsory standards) are established.

In this presentation we will describe experiences and initiatives related to benchmarking to increase and develop the use of such measures.

[1] The project is co-funded by the European Commission DG Information Society and Media, under the contract IST-004526.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 258 – 262, 2006.

2 Soft Measures

These measures may include benchmarking, public procurement, introduction of quality marks, education of people involved in the Internet business, general awareness raising and other measures based on voluntary actions.

In this paper we will go deeper into benchmarking. This includes the experiences from a yearly assessment of Norwegian public web sites and the use of the results to create awareness, plus a short description of EIAO – European Internet Accessibility Observatory – where large scale benchmarking is a key part.

3 Norwegian Benchmarking of Public Content - Norge.no

3.1 Background

The aim of benchmarking is to measure and publish the quality of public websites, and thereby "reward" websites of good quality.

Norge.no has established a set of evaluation criteria to evaluate the quality of public websites in Norway. The criteria are divided into three categories: accessibility, user adaptation and content.

The first evaluation was conducted in 2001 and there has been an annual evaluation since 2003. In 2005, approximately one hour was needed to manually evaluate each page with the 34 checkpoints. When all the evaluations are finished, a report including a username/password is sent to each website authority. If they have comments to the evaluation, they can log on to a website and reply with their feedback and comments. In 2005, 698 websites were evaluated, which gives a total of 23 732 checkpoints. There were only complaints regarding 126 of the checkpoints, or 0.5% of the total evaluation.

To ensure criteria quality and relevance, the evaluation criteria are reassessed every second year. This is necessary because of changing demands and technical evolution. It is also important to allow websites to volunteer their opinions and influence the evaluation criteria to develop their motivation for being involved in the evaluation and satisfying evaluation criteria.

3.2 The System Works

In 2004, only 98 websites satisfied 80% or more in the accessibility category. In 2005, the number was 150. The feedback from the websites indicates that much of the improvement is caused by the awareness created by the benchmarking. By doing so, the "good" websites can act as role models and none of the websites want bad publicity.

3.3 Attention and Publicity

To simplify the presentation of the result the websites are given between one and six stars. The full evaluation is available at the norge.no web page. The results from the benchmark are promoted in several ways:

- Presentation at a conference
- Presentation on the norge.no website
- A quality mark to put on the evaluated website (shown below)

Fig. 1. Norge.no quality mark

In addition to these initiatives the press releases several articles comparing the websites.

3.4 Using the Results

The result from the benchmarking is used to identify which areas that need more attention. The chart below shows that many websites are not satisfactory for checkpoint 1.5 and 1.9. Identifying such problems is important because norge.no can make a special effort to improve information on how to satisfactory follow those guidelines.

3.5 Challenges

The main challenge in producing quality websites seems to be lack of awareness and knowledge. Many of the website authorities are not aware of missing accessibility, and when the results are presented they wonder how to make a satis-factory solution.

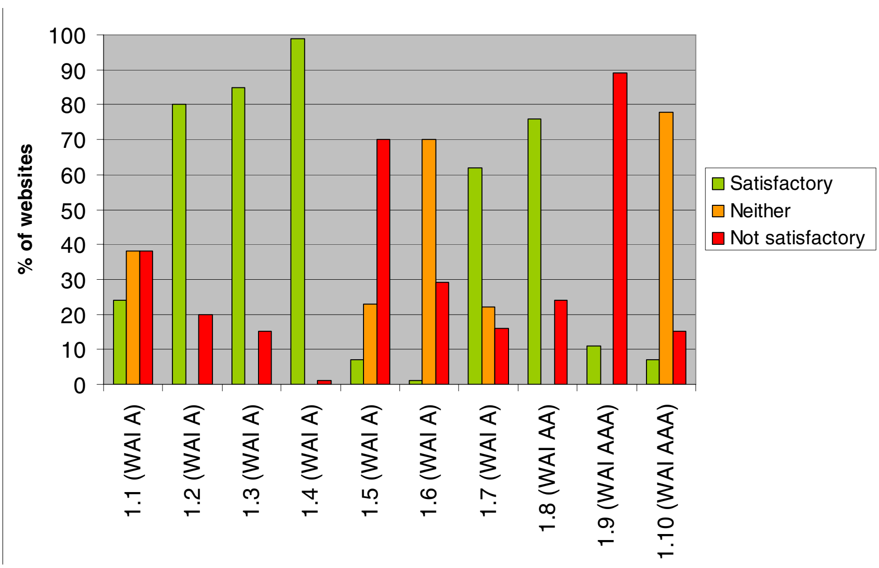


Fig. 2. Kvalitet 2005 (quality 2005), accessibility checkpoints

During the evaluation we get a lot of feedback, and even though we criticize the websites by giving them a bad score, most websites wants to know how they can fix problems and improve the quality, and thereby get better results next year.

4 EIAO

4.1 Background

The overall objective of the EIAO project is to contribute to better e-accessibility for all citizens and to increase use of standards for representing on-line resources.

The project will develop a prototype technical machinery for a European Internet Accessibility Observatory (EIAO) consisting of:

- A set of web accessibility metrics.
- An Internet robot for automatic and frequent assessment of web accessibility and deviations from Web Content Accessibility Guidelines (WCAG).
- A data warehouse providing on-line access to collected accessibility data.

The tools for collecting, assessing and disseminating data will be continuously improved, based on feedback from end users (e.g. policy makers, associations of disabled people etc.) and user tests to sharpen the relevance of the automatically collected data.

In this way, the EIAO can provide large scale benchmarking of websites. Such benchmarking can lead to competition between countries, organisations, developers etc. It may also be used to provide background for political action.

The EIAO Report Web Site will allow for generation of reports for single web sites and NUTS regions *(Nomenclature of territorial units for statistics – NUTS – Statistical Regions of Europe)*. Each of these report types will be available for a given point-in-time as well as for a time period. Consequently, four basic types of web accessibility reports should be available:

- Single web site accessibility scorecard report for a point-in-time
- Single web site accessibility scorecard report for a time period
- NUTS region(s) accessibility scorecard report for a point-in-time
- NUTS region(s) accessibility scorecard report for a time period

Further more a range of criteria according to which the overall accessibility scorecard scores, reported in the four basic types of the web accessibility reports, can be decomposed. These criteria include:

- Disability groups
- NACE branches *(Classification of Economic Activities in the European Community)*
- WCAG checkpoints

The EIAO Report Web Site will allow for generation of user-defined detailed reports where the overall scorecard score from the basic report will be decomposed according to 1, 2, 3, or 4 criteria specified by the user.

Subsequent evaluations of the EIAO Report Web Site will guide future development of the functionality related to the detailed reporting.

4.2 National Accessibility Case

We have established a NAC group related to the project (*National Accessibility Case*) to work with the development of benchmarking combining the experiences from *norge.no* and the prospects of large scale assessment EIAO can offer. Firstly the experiences from using benchmarking as an awareness raisin tool may be used in the development and tuning of the EIAO data warehouse, and secondly the effectiveness of automatic assessment by the EIAO machinery may be used in the future analyses carried out by *norge.no*.

4.3 Conclusions – The Importance of Benchmarking

Even though the sticks often are regarded more important than the carrots we believe the best results will be achieved when using both. Even if the sticks are very effective, using only stick turns the goal of the efforts into a burden rather than a challenge for the people involved. Such reactions may slow down the speed of the desired development. Benchmarking is also a key element in planning activities, as it gives feedback on the effectiveness of the actions carried out to follow up planning. Thus it may contribute to a better basis for political actions and the use of tools in the concerted actions to reach accessibility goals.

References

1. Snaprud M.H., Ulltveit-Moe N., Granmo O.C., Rafoshei-Klev M.E., Wiklund A., Sawicka A..: Quantitative assessment of public web sites Accessibility -Some early results, ETSI, Accessibility for all (2003) http://www.etsi.org/cce/proceedings/7_3.htm
2. Snaprud, M.H., Aslaksen, F.: The European Internet Accessibility Observatory (EIAO) - Aspects of Internationalisation (2004)
3. Craven, J. and Snaprud, M.: Involving Users in the Development of a Web Accessibility Tool. Ariadne Issue 44, 30 July 2005, http://www.ariadne.ac.uk/issue44/craven/intro.html
4. Bühler Ch., Heck H., Perlick O., Snaprud M.: A Framework for Automated Web Accessibility Assessment, AAATE Conference 2005, Lille, 06.09.2005 (2005)
5. Ulltveit-Moe N., Gjøsæter T., Snaprud M.H.: An open architecture for large scale monitoring of web accessibility. (2005)
6. Jansen A., Ølnes S.: Kvalitetsevaluering av offentlege nettstader - Kva er det vi evaluerer? Forvaltningsinformatisk notatserie 3/03, ("Quality evaluation of public web sites" – only in Norwegian) (2003) http://www.afin.uio.no/forskning/notater/3_03.pdf

A Ubiquitous Social Community Portal Service for Social Networking with Convenient Accessibility

Yung Bok Kim

Dept of Computer Engineering, Sejong University
GunJa-Dong, GwangJin-Gu, Seoul, Republic of Korea
yungbkim@sejong.ac.kr

Abstract. A ubiquitous social community portal service for social networking was studied with short International Domain Names (IDN) as a convenient mobile user interface (UI) for the disabled and elderly. Single-character multilingual domain-names are easy for the disabled and elderly people to memorize and type in to access social information. Multi-lingual single-character domain names with text-based social information are more convenient than long URL strings to retrieve information and to give short notice of information for social networking. We introduce a ubiquitous social community portal 'ktrip.net', that has real-time Text to Speech (TTS) functionality with text-based contents in social networking service, using a tiny 'hand-board' for alumni, relatives and any special groups.

1 Introduction

Many elderly people feel social helplessness when coping with age-related changes such as retirement, the death of a spouse, and a diminishing network of surviving peers and relatives, they miss the spontaneous peer contact they enjoyed in previous phases of their lives, but become resigned to solitude because of the tremendous effort required to see friends. Such changes often bring about social disengagement – withdrawal from the relational stimulation and support that can protect against a spectrum of illness [1]. Researchers have also examined social information flows for strong and weak ties in Internet communications and online spaces, how online socializing impacts people's psychological health, and how online socializing affects face-to-face interactions in communities [2]. The rapid modernization through the weakening of place-linked social relations could cause a loss of traditional social community; however we can use ICT to maintain a network of strong social ties with kin, friends, and colleagues who don't necessarily live in the same neighborhoods [3].

We have studied the convenience matters to access to information ubiquitously for social networks, e.g. alumni association, relatives, various offline associations, clubs, and colleagues. Researchers have begun to examine user's finding and refinding behaviors as well as the limitations of existing search technologies [4]. We should extend this concept of finding and refinding behaviors done by the disabled and elderly people to access to information for lifetime social networking service. There are many Internet portals providing various information, e.g. Web page, music, image, video,

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 263 – 270, 2006.

Web site, news, dictionary, encyclopedia, blog, mini-homepage, community café etc. However, the behavior for finding and refinding of right information is difficult or confusing even for skilled users; and the community service cannot be guaranteed for lifetime service until the aged with strong social networking tie. Therefore we need to consider more convenient user interface and simplified portal service adequate in ubiquitous computing and networking environment, especially for the unified and lifetime service until the aged as well as for the offline social networking tie. Social information networking based on Internet, both the wired Internet and mobile Internet, will be made more practical by using short multilingual International Domain Names (IDN), instead of the long and inconvenient URL-string, as a convenient user interface for the disabled and elderly people. We introduce single-character multilingual domain names for Web information access to the ubiquitous information portal in the worldwide multilingual information network based on wired and mobile Internet, and show the useful results from implementation of the ubiquitous portal for social networks using single-character multilingual domain names.

In the following sections, we will discuss multilingual domain names in social networking; here we will introduce the single-character multilingual domain names instead of long URL string for Web information access in social networking. Information access in the Web portal will be discussed with consideration of commercial portals. Then, we will discuss the convenience metric and resource utilization. The empirical results will be discussed on the basis of implementation of information portal, accessible with multilingual single-character domain-names as well as capable of Text to Speech (TTS) functionality for the disabled and elderly people in the Aged Society. Finally we will conclude our study with a consideration of future research.

2 State of the Art for Social Networking

We have studied the usefulness of the International Domain Names (IDN) for information access in social networking, especially the convenience of single-character multilingual domain names; classical bookmark functionality in browser help only limited URLs visited already. For fast and convenient service, speed of typing long and complex URLs is one of the dominating factors in terms of performance for mobile phone users. In the previous work, the performance analysis for user interface in real-time ubiquitous information networks was studied by Kim et al. [5]. Empirical performance analysis of Web accessibility in ubiquitous information networks was also studied [6]. User interface for the disabled and elderly people was discussed in terms of performance in the real-time ubiquitous information network [7]. The focus of the work was on mobile user interface with single-character multilingual domain names. In addition, we studied Text to Speech (TTS) functionality based on speech synthesis in information portals accessible with single-character multilingual domain names in social networking.

The user interface to type in the domain name for access to Internet, the wired Internet as well as the mobile Internet, should be as simple as possible. Single-character multilingual domain names provide a convenient user interface. The scheme for multilingual domain names has been standardized worldwide by IETF (Internet Engineering Task Force) and has been approved by ICANN (Internet Corporation for

Assigned Names and Numbers). A plug-in program for converting from Unicode to ASCII code has been provided by VeriSign Inc., and auto-conversion functionality for standardized multilingual domain name service is embedded in the Web browsers as a built-in functionality, e.g. MS IE7.0, Mozilla Firefox, etc.

For information retrieval and access in social networking, the performance of the unified portal is important to provide QoS (Quality of Service) with cost-effective and inexpensive system solutions. For consistency of information as well as for convenient user interface in social networking, we need a unified portal both for wired Internet and mobile Internet. Fast and convenient access of information as well as notification is required for social networks including mobile application services. We need to write the information or advertisement in the information Web site accessible with single-character multilingual domain names, instead of long URL strings, for convenient Web information access. The writing of information or advertisements on Web sites or special notices for social networking is difficult or impossible, especially for elderly people.

There are many different kinds of portals based on the wired or mobile Internet, and also various community sites are proliferating in online social networks. The complexity of consistent information access from portals or community sites has been increasing, and the inconvenience of the user interface for information access has become serious even for the skilled users. Mobile portals mainly used by the younger generation are more difficult for the elderly because of inconvenient user interface. We suggest a ubiquitous portal based on *'hand-board'* with online white*board* service using *hand*-sized information in PC or *hand*held phone, accessible with simple single-character multilingual domain names as a convenient mobile user interface.

For ubiquitous information portals, unified information service is indispensable in social networking. The ubiquitous Web server should have the capability of showing the unified contents, i.e. the HTML contents for wired Internet as well as the mobile contents for many different kinds of mobile Internet and various contents using different mobile markup languages, e.g. WML (Wireless Markup Language), mHTML (mobile Hypertext Markup Language) etc. Easy typing of the URL was also considered for both wireless and wired Internet. The single-character multilingual domain names, which could become the key character for information access to DB and registration of information as well as advertisement, were considered as a convenient user interface for mobile phones in a ubiquitous information network for the disabled and elderly people in social networking.

3 Convenience Metric and Resource Utilization

For typing the information in a fast and convenient way, the user's typing speed is one of the important performance factors for information retrieval in online social networks, especially with the mobile phone [8,9]. Depending upon the keypad stroke number, the level of convenience of user interface can be estimated as a convenience metric. The convenience metric, defined as the keypad stroke number, helps us to understand the usefulness of single character domain names instead of long URL-strings for unified information service, especially in the mobile Internet environment. For a ubiquitous information network, even the input of characters becomes important for retrieval of information as well as for registration of information, especially with keypads in the mobile phone. To access the unified portal ubiquitously, the user interface for information retrieval should be as convenient as possible for typing in the domain names or URLs for searching the right Web site and information access.

For information access, the text entry with PC or mobile phone is an important basic step to consider the user interface related to the convenience metric. Soukoreff and MacKenzie [10] studied and introduced metrics for text entry, i.e. evaluation of MSD (Minimum String Distance) and KSPC (Keystrokes per Character), and a new unified error metric. This study focused only on the QWERT keyboard. With inconvenient keypads in the mobile phone, the error rate for text entry becomes more serious, and we should keep this in mind for the following discussion.

We can access information in the unified/ubiquitous information portal with multilingual single-character domain names as root nodes in the information tree for social networking. With multilingual single-character .net domains, we can access the required information, especially unified and consistent information in a ubiquitous information network for social networking. If we look at the user interface for handheld phones for mobile information service and even for the URL typing interface for information access in the wired Internet, multilingual single-character is very convenient and useful because it is like a root node in the tree of information-access hierarchy to generate any multilingual domain names or words for information access in China, Japan, Korea, and other countries using multi-lingual domain names.

Even though the single character for Web information access is very convenient, from commercial search portals in service we cannot retrieve any meaningful information even with huge numbers of results after searching with a single character. We need a unified Web information portal, based on the 'hand-board' we introduced, for ubiquitous information network in the ubiquitous networking environment, with convenient accessibility to information for social networking service.

Let's look at mobile user interface in the mobile Internet for ubiquitous information network. The development and application environments are very different from the existing wired Internet environment, i.e. mainly based on MS Explorer as an Internet browser. Most of the Web pages in existence today are designed for desktop PCs, and viewing them on mobile Web browsers is extremely difficult. Chen et al. [11] have studied the adaptation of Web pages for small-screen devices. Adapting Web content to mobile user agents has been studied using an adaptation proxy, XHTML documents into XHTML MP (mobile profile) and WML, by Laakko and Hiltunen [12]. The searched Web information (e.g. mini-homepages, blogs, club, café, etc.) is very difficult to browse with the mobile phone, so we need to consider the ubiquitous portal with convenient user interface and unified (for PC and mobile devices) online bulletin board, (the so called 'hand-board') for text-based information for easy TTS (text to speech) application.

We need to consider the performance of ubiquitous information portals including the Web server as well as the mobile device for considering mobile user interface in the ubiquitous networking environment. We used a single Web server for information access as a unified service for simplicity of management and for cost-effectiveness. This method gives effectiveness and efficiency for the access of information and utilization of resources, in terms of the bandwidth for communication and the size of disk storage to build portals for lifetime social networking service.

We used single Web server 'ktrip.net' for the simplicity of management and the cost-effectiveness of implementation for text-based Web information portal in lifetime social networking service. We considered the effectiveness and efficiency for the management of Web information and utilization of resources for social networking, in terms of the bandwidth for communication and the size of disk storage. Most commercial portals are providing rich-media contents for community, e.g. mini-homepage, blog, club, and café; therefore the resources for DB and Web service are increasing with a lot of investment for computer system and network bandwidth. The size of resources for multimedia service is huge compared to our text-based approach; therefore our text-based portal is self-evidently very cost-effective for lifetime service without further analysis. The finding and refinding of right information in most commercial portals is very difficult. We considered the complexities of consumed resources $\boldsymbol{R}_s$ and $\boldsymbol{R}_p$ (s means pushing scheme, p means polling scheme) for pushing and polling scheme respectively, and the consequent relation between the orders of complexity for resource consumption becomes

$$O(R_s) \gg O(R_p) \tag{1}$$

We need rather efficient way in resource investment with the cost-effective server using polling scheme instead of pushing scheme for online/offline lifetime social networking. This concept will be also important for unified text-based information portal in the mobile Internet environment, for cost-effective ubiquitous communication service.

4 Implementation for Evaluation

Using multilingual single-character domains for fast access to the required multilingual domain name, the required information or advertisement can be registered any time and any place using wired or mobile Internet with multilingual single-character domain names. The implemented system for evaluation is as follows: for the operating system, Windows 2000 server; for wired and mobile Internet Web services, IIS5.0 Web server and ASP 3.0 with MS DBMS; for the mobile markup languages: WML and mHTML for mobile service operators. The Internet communication line to the router and Web server `http://ktrip.net`, i.e. the unified portal for information access in both wired and mobile Internet service, is used on the basis of Metro-Ethernet (10Mbps).

We measured the time to type in the 'ktrip.net' as a simple example to show the critical time in terms of performance. We measured also the aggregate time spent by information server and network, and we observed the time to read the right contents after clicking the title in the displayed information list on the mobile phone screen. We observed that the typing time for full domain names or long URL-string with mobile phone was serious in terms of performance. Therefore we need single-character domain names instead of long URL-strings for mobile Internet applications with mobile phones, as we discussed; and the typing time is faster than 3 seconds, in most cases 1 second.

In terms of packet cost, one packet, i.e. 512 Bytes, costs 0.6 Cents (U.S.), the minimization of delivered packet number from Web server to the mobile phone is important for cost-effective service with ubiquitous information portal. We considered both the cost-effective packet size and the number of packets for delivery in our implementation of ubiquitous information portal for social networking service; and moreover the TTS (text to speech) functionality in service was easy to implement because of the tiny size and text-based 'hand-board' contents.

We did some experimental research for students as follows. The number of students was around 500 during four semesters, i.e. around 125 students in each semester consisting of 4 months, and they used frequently this notification bulletin board so-called 'hand-board' using the site 'ktrip.net'. The cumulative number of usages was around 24,400, and that means the average click number α in one semester by one students was

$$\alpha \approx \frac{24,400}{4 * 125} = 48.8 \tag{2}$$

i.e. around

$$\frac{48.8}{4} = 12.2 \tag{3}$$

clicks per month by each student) thus every student clicked around 3 times in every week. Therefore the suggested method is useful where the usage frequency is around 3 times per week. As an example of Web services for online/offline social networking service, special application for information registration/retrieval can be done with the notification bulletin board (the so-called 'hand-board') for alumni associations in social networks; here the size of the association may be from ten people to a couple of hundred people in online/offline social networking.

Usually alumni association meetings are held between one month and one year apart; that means the information polling frequency by each member is enough to be notified with the frequency between once every two weeks and once every six months. The speed of on-line registration of advertisement/notification as well as the speed of access of special information with information portal is fast enough for real application. Moreover the effectiveness of storage usage for social networking is anticipated if we consider the applications for various communities, mini-homepages, club, and blog services based on rich media, as far as the consumed disk storage and cost for operation and administration of lifetime social networking are concerned.

We implemented also the Text to Speech (TTS) functionality for evaluation, because the simple and text-based information in 'hand-board' service was easy to implement the TTS functionality based on speech synthesis. The conversion time of 1 Kbytes text-based information was around 1 second; and in most cases, the information size for social networking was enough, with around 1 Kbytes. This TTS functionality will be very helpful for the disabled and elderly people; moreover, telematics service for auto drivers will be applicable with convenient user interface for listening to the contents in 'hand-board'.

As we discussed, commercial portals cannot give meaningful information after searching with multilingual single-character or characters, even though it is very convenient to search with single-character as search key for information access in social networking, especially in the mobile Internet using handheld phones with small screens. We are considering the special information service for the disabled and elderly, on the basis of using text-based information portals for social networking services; therefore the listed information, i.e. meaningful information searched based on a multilingual single-character domain name, is not as large as the meaningless information provided by commercial search portal services. We are also considering location-based services using mobile phones, as listed information related to specific locations will decrease tremendously for special applications in the ubiquitous information networking environment. The convenience and usefulness of multilingual single-character will be remarkable for ubiquitous information networking in social networking services, with any mobile devices for the disabled and elderly people in the Aged Society, and without Digital Divide as a lifetime service.

5 Concluding Remarks

A ubiquitous community portal service using multilingual single-character domain names instead of long URL-strings for convenient information access has been studied for lifetime social networking service for the disabled and elderly. The convenience of multilingual single-character domain-names and the tiny 'hand-board' for mobile Internet service was discussed considering state of the art for social networking. The performance for information access using multilingual single-character domain names with mobile phone for mobile Internet was also discussed, considering the keypad-press number with mobile phone as a convenience metric. The results of implementation show the usefulness and convenience of multilingual single-character domain names for information access in social networking services. Additionally, the Text to Speech (TTS) functionality based on the text-based contents in 'hand-board' service for social networking was implemented in the unified site 'ktrip.net' to increase the convenience in the ubiquitous information portal for the disabled and elderly people in the Aged Society. With a more ubiquitous network environment, the delay time related to user interface will become more critical; thus we need more efficient user interface for lifetime social networking, especially using mobile Internet with mobile devices. For further research, the speech recognition functionality for the disabled and elderly people in the Aged Society will be studied to increase the convenience metric as well as to decrease the delay time, i.e. performance metric, for search and notification in lifetime social networking services, considering multimedia information access.

References

1. Morris, M.E.: Social Networks as Health Feedback Displays. IEEE Internet Computing, pp. 29-37, September/October, 2005
2. Churchill, E.F. and Halverson, C.A.: Social Network and Social Networking. IEEE Internet Computing, pp. 14-19, September/October, 2005
3. Jones, Q. and Grandhi, S.A.: P3 Systems: Putting the Place Back into Social Networks. IEEE Internet Computing, pp. 38-46, September/October, 2005
4. Capra III, R.G. and Perez-Quinones, M.A.: Using Web Search Engines to Find and Refind Information. IEEE Computer, pp. 36-42, October, 2005
5. Kim, Y.B., Kwak, M. and Cho, D.: Performance Analysis for User Interface in Real-Time Ubiquitous Information Network. APCHI 2004, LNCS 3101, pp. 171-179, 2004

6. Kim, Y.B. and Kim, Y.H.: Empirical Performance Analysis of Web Accessibility in Ubiquitous Information Network. UI4All 2004, LNCS 3196, pp. 268-274, 2004
7. Kim, Y.B. and Kim, J.Y.: Performance Analysis of User Interface for the Disabled in Real-Time Ubiquitous Information Network. ICCHP 2004, LNCS 3118, pp. 926-929, 2004
8. Kim, Y.B. and Cho, D.: Management of Korean Alphabet Domain Names for Ubiquitous Korean Information Network. UKC 2005, SanDiego, USA, August 11-13, 2005
9. Kim, Y.B. and Kim, Y.H.: King Sejong's Invention: Korean Alphabet and Mobile User Interface for Information Society. International Conference, UI4All, Vienna, Austria, June 2004
10. Soukoreff, R.W. and Mackenzie, I.S.: Metrics for text entry research: An evaluation of MSD and KSPC, and a new unified error metric. CHI 2003, Ft. Lauderdale, Florida, USA, April 5-10, 2003
11. Chen, Y., Xie, X., Ma, W.Y. and Zhang, H.J.: Adapting Web Pages for Small-Screen Devices. IEEE Internet Computing, pp.50-56, January/February, 2005
12. Laakko, T. and Hiltunen, T.: Adapting Web Content to Mobile User Agents. IEEE Internet Computing, pp. 46-53, March/April, 2005

Web Accessibility Conformity Assessment – Implementation Alternatives for a Quality Mark in Austria

Marie-Luise Leitner[1], Klaus Miesenberger[2], Daniela Ortner[2], and Christine Strauss[3]

[1] University of Vienna, Department of Finance, Bruenner Str. 72, 1210 Vienna, Austria
marie-luise.leitner@univie.ac.at
[2] University of Linz, Institute Integriert Studieren, Altenberger Str. 69, 4040 Linz, Austria
{klaus.miesenberger, daniela.ortner}@jku.at
[3] University of Vienna, Department of Business Studies, Bruenner Str. 72, 1210 Vienna, Austria
christine.strauss@univie.ac.at

Abstract. Various European Union initiatives have focused on the dissemination and harmonization of approaches for assessing the conformity of web accessibility. This paper suggests a scenario-based decision support for the implementation of a web accessibility quality mark in Austria on the basis of a framework proposed in a CEN Workshop Agreement. The paper analyzes different implementation alternatives in order to facilitate and accelerate the realisation of such approaches at the national level and to encourage other European countries to adopt selected elements for their own initiation.

1 Introduction

In recent years, web accessibility quality marks have been developed and established on a national basis in several European countries. Although they are all based on the Web Content Accessibility Guidelines 1.0 [18] published by the World Wide Web Consortium, a variety of different evaluation procedures, implementations and levels of conformity have led to considerable heterogeneity within the European context. To avoid further fragmentation, the "Support-EAM" project (Supporting the creation of an e-Accessibility Mark) [16] was initiated by the European Commission as part of the eEurope 2005 Action Plan [4]. One of Support-EAM's objectives involved the development and stipulation of requirements for a harmonized European web accessibility conformity assessment scheme, which was realized within a CEN (European Committee for Standardisation) Workshop. As a result of this cooperative effort between industry and consumer organisations, the CEN Workshop Agreement proposed specifications for a European web accessibility conformity assessment scheme and the relevant quality mark.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 271–278, 2006.

This paper is based on that CEN Workshop Agreement [3] and explores viable alternatives for implementing the European web accessibility quality mark in Austria. This article applies a look-ahead approach that assumes the release of a normative document and an evaluation methodology in the near future. The scenario analysis includes the development of four scenarios and their evaluation in terms of six criteria.

2 State of the Art in Web Accessibility Conformity Assessment

2.1 Benefit of Web Accessibility

Realizing web accessibility has a positive impact on a company's reputation and image [14]. The risk of possible litigation costs that could occur if a website is not fully accessible is eliminated; in addition, a fully accessible website enjoys a larger potential audience and customer base. A rough cost-benefit analysis that takes into account the total accessibility costs, which depend on the enterprise size and the complexity of the website on the one hand and the increased audience size on the other, suggests that accessible websites enjoy estimated relative savings of 12-35% of the website costs [11]. Validated websites could indicate the accessible character of a particular purchasing process, to name just one example, for customers with special needs. A harmonized European web accessibility quality mark communicates these efforts and at the same time fosters the promotion of web accessibility in Europe [10].

2.2 Legal Framework in Austria

The rights of people with disabilities are protected by law in an increasing number of European countries. In Austria, Article 7 of the Austrian Federal Constitution of 1997 states that nobody should be discriminated against due to an impairment [12]. In addition, the Austrian Equalization Act for People with Disabilities establishes that people with impairments must be granted equal rights for participating in public life [1]. Some European countries have enacted laws to oblige (public) websites to be accessible. The Austrian E-Government Act of 2004 stipulates that public websites have to meet international standards on web accessibility by January 2008 [13]. Compliance with such legal obligations could be verified with the help of a quality mark, which would also strengthen public awareness for the issue of web accessibility and thus make an important contribution to the objectives of the eEurope 2005 Action Plan [4].

2.3 European Web Accessibility Conformity Assessment Scheme

The CEN Workshop Agreement establishes the European Authority for Web Accessibility Conformity Assessment (EAWAC) as the central European body with ownership of the relevant quality mark. The creation and operation of a European conformity assessment scheme for web accessibility is one of EAWAC's main responsibilities. This tasking includes the drafting of a normative document to set criteria for web accessibility, as well as the establishment and administration of a quality mark.

Different options exist for the framework in which the mark is implemented at the national level; however, each of these assumes the existence of one unique normative document created by the EAWAC. The CEN Workshop Agreement does not recommend or enforce the implementation of a distinct framework option, nor does it determine the number of options being implemented on a national level.

1. *Inspection:* The ISO/IEC 17020 [6] accredited inspection body issues the quality mark. Furthermore, the inspection body performs regular surveillance of the certified websites and withdraws the quality mark if the criteria have not been met. Membership in the EAWAC is mandatory for inspection bodies.
2. *Product certification:* In this option, a EN 45011 [5] accredited product certification body issues the quality mark. The product certification body performs regular surveillance of the certified websites and withdraws the quality mark in case of non-compliance with the criteria. The product certification body must be a EAWAC member.
3. *Supplier's declaration of conformity:* The owner of a website assures that his/her website meets the criteria set in the normative document issued by the EAWAC. The owner accepts the specifications presented in the ISO/IEC standard 17050 part one [7] and two [8] about the supplier's declaration of conformity. The website owner receives the quality mark after registering with the EAWAC, which performs regular surveillance to assure compliance.

3 Implementation Alternatives for a Web Accessibility Quality Mark in Austria

3.1 Conditions

Discussions with national stakeholders in Austria about the CEN Workshop Agreement led to a serious debate that identified some weaknesses that the scheme had with regard to its implementation in the country.

Normative Document. A European quality mark should be based on a normative document that stipulates the specifications for certification. With regard to web accessibility, the requirements will include the Web Content Accessibility Guidelines for accessible web presence. An evaluation methodology that contains a procedure for testing the fulfilment of the requirements is essential for the certification process. At present, the EU Web Accessibility Benchmarking (WAB) Cluster [9] is working on a Unified Web Evaluation Methodology (UWEM) [17] that could be used as a basis for that purpose. National stakeholders have pointed out that the development of these normative documents is a prerequisite for the implementation process. These documents can take the form of either a European Norm or a CEN Workshop Agreement.

Certification Types. The CEN Workshop Agreement scheme proposes three alternative mechanisms for implementing web accessibility conformity assessment: (i) inspection, (ii) product certification, and (iii) supplier's declaration of conformity (cf. Section 2.3). The major difference between inspection and product certification does not lie in the certification process, but rather in how the object is defined. In the approach outlined in this paper, no distinction is made between inspection and product certification, as both alternatives represent third party systems with similar structures.

Due to cost aspects, a simultaneous implementation of third party approach and supplier's declaration of conformity would make the third party approach obsolete. For this reason, national certification bodies might be less inclined to cooperate if the supplier's declaration of conformity were to be introduced in parallel.

3.2 Evaluation Criteria

European progress in establishing web accessibility conformity assessment suggests a variety of possible scenarios, whose individual applicability can be compared and analyzed by means of six evaluation criteria: complexity, costs, dependence, flexibility, impartiality and time.

1. *Complexity.* The complexity of implementation depends on the quality mark's background structures. This criterion represents a measure for the amount of prerequisites necessary for implementing the scenario. Thus, a scenario's complexity increases with the existence of ownership and license agreements at a national or at the European level.
2. *Costs.* A scenario's costs comprise the setup costs of the national issuing organisation, the issuing costs of the quality mark or certificate (accreditation and testing costs) and the licence costs for the quality mark imposed by the EAWAC. These costs, especially the setup and testing costs, cannot be specified as they depend on a number of variables, such as the providers organisational form (profit or non-profit), the sample size or the evaluation procedure. All of these variables must be specified in the normative document and the evaluation methodology; as a result, they are not known at present. This paper compares the scenarios in terms of the existence of various cost elements.
3. *Dependence.* Some scenarios can only be realized if certain prerequisites have been fulfilled. These may depend on national, European and international certification, accreditation or legislation bodies. This criterion expresses the degree to which an implementation scenario is dependent on further authorities or institutions. The dependence on the normative document and the evaluation methodology holds true for every scenario and, subsequently, does not need to be taken into consideration.
4. *Flexibility.* The flexibility of implementation represents the issuing organisation's degree of self-determination. For all of the scenarios presented in this paper, the quality mark must conform to a standard or a workshop

agreement. However, the administrative background and, therefore, the degree of flexibility vary from scenario to scenario.

5. *Impartiality.* The quality mark system's impartiality is measured by this criterion, which compares third and first party conformity assessment in terms of their objectivity.
6. *Time.* The time period from the development of the normative document to the implementation of a particular scenario is covered by this criterion.

3.3 Implementation Scenarios

This section outlines four basic scenarios and analyzes them by means of six evaluation criteria (cf. Section 3.2) with the aim of supporting and accelerating the national implementation process once a European normative document or a CEN Workshop Agreement has been released.

Scenario 1: Supplier's Declaration of Conformity. Scenario one includes a first-party evaluation by a supplier according to the international standard ISO/IEC 17050. This approach is easy and cost-effective to implement, as no accredited third party is involved in this scenario. The owner of a website can declare his/her commitment to a normative document or a CEN Workshop Agreement. He/she is authorised to place the quality mark on the website (for a limited period, e.g. for one year), provided that the relevant criteria have been fulfilled and the supplier's intention to use the quality mark has been communicated to the EAWAC [3]. However, if the criteria have not been met, website users may post complaints and the EAWAC can impose sanctions or withdraw the right to issue the quality mark in the event of violations.

Scenario one constitutes the least complex scenario, as it does not require third party involvement (as a consequence, no accreditation procedure accrues). For the same reasons, scenario one is an especially cost-effective and flexible solution. However, it is also highly dependent on existing structures, as it can only be established once the European Authority and a quality mark have been created. The impartiality of this scenario is a matter of great discussion. A study of e-commerce and financial websites indicated that 30% of the websites that had an accessibility quality mark claimed a higher level of accessibility than they actually provided [15]. Because the quality marks concerned had been issued with a supplier's declaration of conformity, a certain lack of impartiality should be assumed for the first party system.

Scenario 2: Product Certification without a Quality Mark. In the second scenario, an Austrian certification body would issue an explicit certificate stating conformity with the normative document or the CEN Workshop Agreement. Scenario two could provide a temporary alternative until the EAWAC, as the owner of the European quality mark, and the corresponding structures are established. This scenario requires neither the elaborate structure of a European quality mark nor ownership or license agreements.

Accordingly, scenario two is a third party solution with relatively low costs and - compared to scenario three - with high flexibility. The low complexity of this scenario is attributable to the fact that it is independent of administrative structures. The impartiality of third party conformity assessment makes scenario two a competitive solution that can be implemented immediately after the release of the normative document.

Scenario 3: Product Certification with a Quality Mark. Scenario three assumes both the release and approval of a European quality mark and the establishment of an organisational structure for a specific web accessibility label. Thus, scenario three represents a follow-up or stage of expansion to scenario two. The quality mark in scenario three would be issued by national certification or inspection bodies.

This scenario involves the most elaborate structures. The EAWAC and the European quality mark system must be set up before scenario three can be realized, making the scenario highly dependent on European authorities and structures and therefore inflexible in its implementation. The costs for scenario three exceed the costs for scenario two, due to the additional licence costs incurred for the European quality mark. Because it results in the issuance of a harmonized European quality mark, scenario three, like scenario one, is in keeping with the final goal of the CEN Workshop Agreement.

Scenario 4: Product Certification Using Existing Structures. Scenario four relies on well-established organisational structures and quality marks, such as the Keymark. The Keymark is an existing quality mark that stands for the compliance of products with requirements of the relevant standards. It has developed as an umbrella label in the technical sector and currently consolidates 150 European standards [2].

If the Keymark were to become the operating mark, the EAWAC would need to be embedded into CEN, the owner of the Keymark. Such a scenario goes beyond the scope of the CEN Workshop Agreement, which does not consider the use of existing structures for the creation of a quality mark. As the issuing organisation has already been established, the costs, particularly the setup costs, may be lower than in scenario three, provided that the owner of the existing mark refrains from any additional fees for the structure and label. This factor makes the implementation of scenario four highly dependent on the authority owning the existing mark. Adopting an existing structure assumes that the owner and issuer fully agree. In turn, such an agreement might require negotiations that could delay and hamper the implementation of scenario four. Existing structures may decrease the complexity of implementation, but they also keep the flexibility to a minimum. A main advantage of this scenario lies in its not creating any additional administrative and bureaucratic structures. Its impartiality is given through third party certification.

3.4 Scenario Evaluation Overview

The rough evaluation of the four scenarios by means of six criteria (cf. Section 3.2 and 3.3) provides a basis of support for decision-making on the national

Table 1. Scenario Evaluation

Criteria	Scenario 1	Scenario 2	Scenario 3	Scenario 4
Complexity	low	low	high	medium
Costs	low	medium	high	medium
setup	*no*	*yes*	*yes*	*no*
issuing	*no*	*yes*	*yes*	*yes*
licence	*yes*	*no*	*yes*	*yes*
Dependence	high	low	high	very high
Flexibility	high	medium	low	very low
Impartiality	low	high	high	high
Time	sooner	immediately	sooner	later

implementation of the quality mark. Table 1 gives an overview of the four alternative scenarios, taking into account the criteria complexity, costs, dependence, flexibility, impartiality and time. In an early stage of realisation, scenario two is a reasonable strategy, as no explicit quality mark system is needed. Once the EAWAC and the quality mark are established, a follow-up choice can be made between scenario one, as a cost effective solution with a lack of impartiality, or scenario three, whose third party conformity assessment brings with it higher costs for the end user. Scenario four may require negotiations with existing quality mark owners, but could result in a cheaper and less complex alternative.

4 Summary and Perspective

Harmonization of web accessibility conformity assessment has become an issue of growing importance for the European Union. This contribution uses a scenario technique in order to provide initial decision support for the implementation of a web accessibility quality mark in Austria. The purpose of this paper is to analyze the advantages and drawbacks of various implementation alternatives in order to facilitate and accelerate their national realization. Additionally, the requirements for each scenario are examined and categorised in terms of their complexity, costs, dependence, flexibility, impartiality, and implementation time. As a next step, quantitative methods should be applied in a refined analysis in order to develop a business plan and provide decision makers with detailed cost and benefit estimates. This contribution intends to stimulate authorities in other European countries to consider the proposed scenarios from their own points of view and to adopt those elements that seem appropriate for their national implementation.

References

1. Bundesgesetz über die Gleichstellung von Menschen mit Behinderungen (Austrian Equalization Act for People with Disabilities), August 2005, http://portal.wko.at/wk/dok_detail_file.wk?AngID=1&DocID=381322&DstID=252&StID=193960

2. CEN/CENELEC Internal Regulations, Part 4: Certification: The CEN/CENELEC European Mark System, Second Revision, July 2001. http://www.cenorm.be/BOSS/supporting/reference+documents/ir4e.pdf
3. CEN Workshop Agreement on Specifications for a Complete European Web Accessibility Conformity Assessment Scheme and a Web Accessibility Quality Mark, March 2006. http://www.cenorm.be/cenorm/businessdomains/businessdomains/isss/activity/wswacfinalcwa.pdf
4. Communication from the Commission to the Council, the European Parliament, the Economic and Social Committee and the Committee of the Regions: eEurope 2005: An information society for all, An Action Plan to be presented in view of the Sevilla European Council, 21/22 June 2002, http://europa.eu.int/information_society/eeurope/2002/news_library/documents/eeurope2005/eeurope2005_en.pdf
5. EN 45011:1998 General requirements for bodies operating product certification systems (ISO/IEC Guide 65:1996)
6. EN ISO/IEC 17020:2004 General criteria fort the operation of various types of bodies performing inspection
7. EN ISO/IEC 17050-1:2004 Conformity assessment - Supplier's declaration of conformity - Part 1: General requirements
8. EN ISO/IEC 17050-2:2004 Conformity assessment - Supplier's declaration of conformity - Part 2: Supporting documentation
9. EU Web Accessibility Benchmarking (WAB) Cluster Homepage, http://www.wabcluster.org
10. European Design for All e-Accessibility Network (EDeAN) White Paper: Promoting Design for All and e-Accessibility in Europe, 2005.
11. Heerdt, V., Strauss, C.: A Cost-Benefit Approach for Accessible Web Presence, in: Klaus J., Miesenberger K., Burger D., Zagler W. (eds.) Computers Helping People with Special Needs, Lecture Notes in Computer Science (LNCS), Vol. 3118, Springer, Berlin 2004, p. 323-330.
12. Österreichische Bundesverfassung (Austrian Federal Constitution), 2000, http://www.austria.gv.at/Docs/2005/8/30/bundesverfassung.pdf
13. Österreichisches E-Government Gesetz (Austrian E-Government Act), Oct. 2004, http://www.cio.gv.at/egovernment/law/E-Gov_Act_endg_engl_Fassung1.pdf
14. Ortner, D., Miesenberger, K.: Improving Web Accessibility by providing higher education facilities for Web designers and Web developers following the Design for all approach, in: Tjoa, A M., Wagner, R.R. (eds.): DEXA 2005; Database and Expert Systems Applications, Proceedings of the Sixteenth International Workshop, Copenhagen, Denmark, IEEE Computer Society, Brussels, p. 866-870.
15. Petrie, H., Badani, A., Bhalla, A.: Sex, lies and web accessibility: the use of accessibility logos and statements on e-commerce and financial websites, presented at the Accessible Design in the Digital World Conference, Dundee, Scotland, 2005. http://ewic.bcs.org/conferences/2005/accessible/session5/paper2.htm
16. Support-EAM Project Homepage, http://www.support-eam.org
17. Unified Web Evaluation Methodology 0.5, http://www.wabcluster.org/uwem05/
18. World Wide Web Consortium, Web Content Accessibility Guidelines 1.0, http://www.w3.org/TR/WCAG10

Accessible Interfaces to Empower the User

Mari Luz Guenaga, Javier Oliver, and Ander Barbier

Faculty of Engineering, University of Deusto,
Avda. de las Universidades, 24. 48007 Bilbao, Spain
{mlguenag, oliver, barbier}@eside.deusto.es
http://www.eside.deusto.es

Abstract. Highly interactive interfaces, where complex information is presented and managed, need further research. Existing guidelines, techniques and recommendations focus on the user accessing, reading and understanding content and services, but there is the need to go beyond, and convert users into authors, editors or managers of products and services provided by ICT. DAIA project (*Accessible Design of Advance Interfaces, a practical application*) has contributed to the understanding of human behavior in this context and to a better approach of guidelines to improve this kind of interfaces. Users with disabilities have broadly contributed to the success of this project through a web questionnaire and a software prototype testing.

1 Introduction and Aim

Guidelines, techniques and recommendations developed by organizations enable users with disabilities to access products, services and content provided through telecommunication and information technologies (ICT). Current research and development focus on web content, user agents and authoring tools, what makes it easier for users to benefit from the information society.

Even though there is a long way to walk in these areas, many improvements have been done on the accessibility and usability of interfaces, and it is time to step over this point of view of accessibility, empowering the user, what means that we have to provide usable highly-interactive interfaces, so they are able to create, edit or manage content, services and products.

READIS (*Accessible Resource Centre for People with Disabilities*) project first [1], and DAIA at the moment, aim to improve usability and accessibility of user interfaces which require intense interaction with users and where complex data is managed. With this purpose a web questionnaire and two software prototypes have been developed, to gain applicable knowledge about specific needs and preferences of people with disabilities facing these kind of interfaces.

The outcome of users' response via web questionnaire provided input to the design and engineering process which aim was the specification and implementation of improved prototypes.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 279 – 286, 2006.

2 Methods

READIS and DAIA projects have involved users from the very beginning of the design process. End-user organizations and individuals have been contacted to participate via Internet questionnaire or personally in prototype testing.

2.1 Questionnaire, Quantitative Research

After the review of existing accessibility guidelines and techniques we have used a questionnaire, via Internet, to obtain information about problems users still face accessing user agents, web sites and specifically when using web forms and managing hierarchical information.

The pre-questionnaire design, first in paper and later in web format, was evaluated by experts and users so it conformed the following requirements:

- Easy and fast to answer
- Comprehensible questions and options to answer
- Usable and accessible interface

The preliminary test give way to the spread of the existence of the questionnaire, mainly through email, so Spanish-talking users could answer it. The questionnaire was four months on the Web, from July to October 2005.

A preview of users' answers to the questionnaire, and their receptive attitude, favored individual contact with some of them to obtain detailed information about their answers.

2.2 Prototype Testing, Qualitative Observation

Volunteers, chosen from the group of people who answered on-line questionnaire, where asked to test two software prototypes of a digital resource centre (from now on DRC). The knowledge about user preferences and problems, obtained from their answers, and the study of usability engineering lifecycle, guidelines, methodologies and techniques [2,5,6] led us to develop two prototypes. Both of them have some elements in common, for example:

- Interface structure and general design. The number, meaning and initial position of content blocks or sections is the same in both prototypes (see Fig. 1).
- User interface configuration options. These options make it possible to personalize elements like font, foreground and background colors and styles, and what is more significant, users are able to hide or show the sections of the interface (heading, path, author's bar, keyboard shortcut bar, etc.), and they are also able to place these blocks in the preferred position: up or down, right or left (what means before or later for users with screen readers).
- Content and basic functionality to perform the test. Only the necessary parts of the DRC were implemented to carry out the test, and in both cases the categories, resources, attributes and content in general was the same.

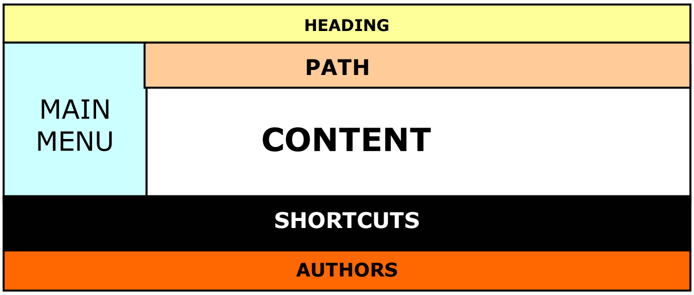


Fig. 1. Interface structure and initial block position

These two prototypes also have significant differences, which aim to contrast interaction mechanism and preferences:

- Main menu structure, guided by concepts-actions in one of them, and in a concept guided way in the other one.
- One of the prototypes has more help and assistance than the other (*see Fig. 2*), in the way assistant programs do.

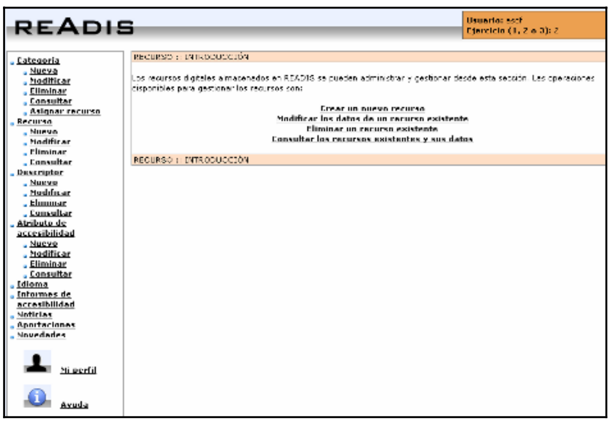


Fig. 2. Screenshot of one prototype, the one that follows a more guided model

The aim of the test was:

- To observe and contrast user's behavior and problems using two different interaction mechanisms.
- Users have to play the role of content administrators of the DRC, so they have to interact in depth with the interface and manage complex data.

The steps followed before the very test started were the following [3]:

1. Contact users and ask for their availability for a 2.5 hour test.
2. Schedule the appointments with all users, grouping them by place of residence.
3. One week before the test we sent them the user guide of both prototypes.

4. Before the test we asked the user personal information and data about their computer, web browser and Internet connection, in order to foresee problems due to configuration or hardware performance.

Once we were in the place of the test, the procedure was as follows:

1. Explained the user the aim of the test, the way the DRC worked and the main differences between both prototypes. We also answered all questions he or she may pose.
2. We presented the prototype that provides more assistance for a beginner, and asked for initial configuration options.
3. The user logged in to complete the first task. Each user was intended to complete three tasks with each prototype, launched in order of difficulty, but if any problem to fulfill the task came up we presented only one or two tasks. Therefore, the total length of the test, for the welfare of the user, was no more than two hours.
4. During the task the evaluators took note of the user's behavior, difficulties, way of browsing the web or interacting with the system, besides all comments and opinions. Prototypes themselves also recorded all user's interaction in a log file.
5. The task was considered finished if the user got the objectives, if too much time was spent or if the user found any problem that made it very difficult or impossible to finish the task in a reasonable time. The minutes spent to finish each task were noted, not as an accurate data but as additional information.
6. After finishing the tasks with the first prototype, steps 2 to 5 were repeated for the second one.
7. Once they had finished all tasks, and with a better knowledge of the system, users were asked to re-select configuration options given by the system.
8. They were also asked about some specific elements like the best way to presentation hierarchical information, accessibility attributes of digital resources, structure of the main menu and interaction mechanisms, or the best way to manage elements of the system (categories, resources, attributes or keywords).
9. As a conclusion, they should compare both prototypes, arguing the best and worst properties of each, and their advantages and disadvantages.

3 Results of Prototype Testing

User testing was made with nine blind people who used JAWS screen readers [4], and 44% of them also used Braille displays. The average was 44 years old and 13.6 years using ICT. 40% were female and 60% were male.

3.1 Observing the Users

First of all we will describe the conclusions achieved from the observation of users facing tasks with both prototypes, how they interact with the system, problems arisen during the test and their general behavior and opinion.

Navigation and User Profiles. All users start the test browsing the system sequentially at least once, looking for the general structure of the web, the different sections, content and the way it works. Once they are familiar with the site, some of them use more advanced techniques to scan it faster, for example they use searching facilities and lists given by the screen reader (link, section or frame lists).

45% of users browse the web exclusively in sequence, what means that sections and important content must be clearly identified with meaningful names, and useless or repeated content must be removed, so users can go as straight as possible to the content. Otherwise users tend to accelerate the processing of the page, missing essential information.

The ability and knowledge of users handling screen readers make a big difference in user's profile. As well as using advanced techniques described before, they are able to use not usual controls, needed in interfaces for managing complex information, like simple or multiple selection lists, combo boxes, pull-down lists, etc. Beginners found them very difficult or impossible to use.

Comparing Prototypes. One of the problems at the beginning of the test was the lack of knowledge on how the DRC worked and the concepts in which it relies. Although users had the user's guide before the test, most of them had not read it in depth, so explanations and help was necessary before and during the test.

The preferred prototype for the 67% of users was the named *Prototype 1*, the one with no so much help, in which resources and content was managed directly accessing it (*see Fig. 3*). The reasons given by users are the more direct access to information, the fewer number of links and the simplicity of the design and structure of the content, as they say "*it is more natural*". All these properties make it easier and faster to complete the tasks using this prototype.

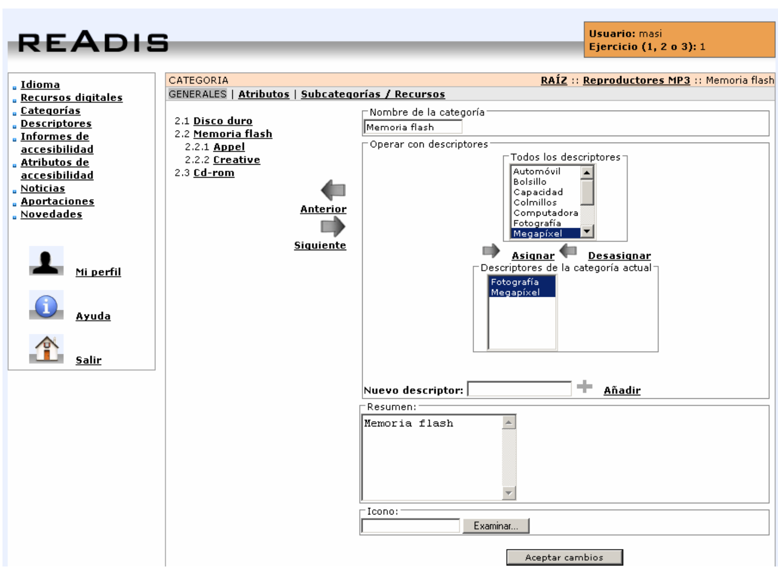


Fig. 3. Screenshot of the preferred prototype for the 67% of users

3.2 Interface Configuration

At the beginning of the test users have the opportunity to select some options to personalize the interface, as they were not familiar with it, 80% of them selected the predefined blind profile.

At the end 90% of users preferred to show only text if it is possible and not icons or images at all. As far as block configuration concerns, users' choice is shown in table 1:

Table 1. Users' choice for the configuration of content blocks, those shown in figure 2. Grey areas are options not possible to choose. The value is the percentage of users who chose that options. The percentage of position options (up, down, both, left, right) are over the number of users who chose to make them *visible.*

	Visible	Invisible	Up	Down	Both	Left	Right
Heading	45%	**55%**	100%	0%			
Path	**67%**	33%	67%	33%	**0%**		
Main menu						11%	67%*
Keyb. Shortcuts	**67%**	33%	17%	83%			
Authors' bar	33%	**67%**	33%	67%			

* For the remaining 22%, it is the same to place the menu on the left or on the right.

The trend in general, and for beginners in particular, is to clear the interface as much as possible, even more if the system is a daily-use tool: 55% of users chose to hide the heading and 67% of them chose to hide the authors'/copyright bar.

The reader may ask why this percentage is not bigger, or why the 67% chose to make *path* and *keyboard shortcuts* visible. In the case of the *path*, users find it more useful, because it provides context information, and in the case of keyboard shortcuts or the selection of other blocks as visible (*authors' bar* or *heading*), it is because they serve as references to identify the beginning and end of content blocks, what should be done in another, more appropriate way.

It is also important to stress that none of the users asked want to repeat path information, or any information at all. Repeating information is very confusing, even designers may think it helps them to understand or remark important information, because they get lost in the structure and content of the web site.

These results show that non essential or repeated elements should be removed from the interface, or at least give the opportunity to do so, and content should be clearly structured and identified.

The last question about user configuration was about how to show hierarchical information, 56% percent selected a numbered schema-like list, better than a dotted item list.

3.3 Specific Interface Elements

Users were asked to give a value between 1 (not satisfying at all) and 7 (very satisfying), and their opinion about some specific elements of the user interface, that we considered specially important. The number obtained, shown with each item into brackets, is very significant because the number of users is low, but it is a hint about their preferences.

Keyboard Shortcuts (4.78 average mark (am)). Users find them useful if the web is a tool to be used daily. But the lack of a standard set of keyboard shortcuts and the occasional use of many pages makes it not worth to memorize them. In addition, sometimes these keyboard shortcuts overlap browser's or screen reader's shortcuts, disabling their more important functionality.

Hidden Links or Shortcuts (4.17 am). Hidden links to skip blocks of content or to access directly to specific sections are worse valued. This is due to the fact that they add text and content to the web page, and many times users do not even realize what they are for. Their advantage compared to keyboard shortcuts is that they do not need to be memorized, and as users get familiar with them their opinion is better.

Configuration Options (5.78 am). This is the best valued item of the test, because it enables to make the web lighter and adapt it to their preferences, so the DRC becomes more usable.

Hierarchical Information Tree (5.44 am). Categories are shown like a two-level-depth tree, where each entry is a link that expands, if selected, in subcategories and resources. Users agree that two levels is enough, 22% of them even think that expanding only the selected category would be enough.

The preferred way of handling hierarchical information is the way Microsoft Explorer does, expanding, contracting or scanning items with cursor keys. The fact is that this element is not available for web sites, there are visually similar controls but they do not work properly with screen readers.

Simple Selection List (3.22 am). This has been one of the most problematic elements of the interface. In some cases it has limited the number of tasks proposed to the users, because the correct use of lists was essential to complete the task.

The problem lies on the difficulty of using this control, it is not a common element on average web sites, designed to be browsed and read by users. Yet, it is frequently used in administrating or managing sites. Only advanced users were able to enter the list, select the desired item and leave it, without loosing or blocking the web site. Other users prefer an alternative way to select items, typing if possible.

4 Conclusions

All users implied in the test of a digital resource centre were blind people using screen readers, in some cases they also used Braille displays. The aim of the test was to gain a better understanding of their behavior, interaction mechanisms and preferences using an interface, where complex data is shown and high interaction is required to complete the management tasks.

Better skills using assistive technologies, and ICT in general, make a big difference in user's satisfaction, but this should be overcome making more usable interfaces. No matter what their experience was, all of them make a sequential reading of the web page at least once, and for many of them this was the only way to browse a site. That is why interfaces must be as simple and clear as possible, removing useless and repetitive information, so the user goes straight to the content.

Configuration options were very valuable, as users could remove or replace blocks of content in the preferred side of the web page. In addition, sections and important information must be clearly identified, and repetition of context information should be avoided, this may lead users to loose themselves in the page or to miss important information.

General guidelines apply in advance and complex interfaces, but there is the need to stress in simplicity and the existence of alternative ways to interact with the system.

There is a hard work to improve the usability of interfaces where users have to carry out management and administration tasks. In this sense, it is essential to provide more and more accessible and usable services and tools where users are not only the customers or users, but the service providers and managers.

Acknowledgement

We would like to thank all users that have answered the on-line questionnaire, responded to our requests for more information and participated in tests and interviews. Thank you for your kindness and availability. Without them this work would not have been possible.

References

1. Barbier, A.; Guenaga, M.L.; Eguíluz, A.; Oliver, J.; Jamardo, I. (2004) READIS - Centro de recursos digitales accessible. Actas de las I Jornadas de Accesibilidad y Nuevas Tecnologías – JANT 2004
2. Good, M., Spine, T. M., Whiteside, J. and George, P. User derived impact analysis as a tool for usability engineering. In Proc. CHI '86 Human Factors in Computing Systems (Boston, April 13-17, 1986), ACM, New York, pp. 241-246
3. Pernice, K and Nielsen, J. (2001). How to conduct usability evaluations for accessibility: Methodology guidelines for testing websites and intranets with users who use assistive technology. `http://www.Nngroup.com/reports/accessibility/testing`
4. Nielsen, Jakob, and Landauer, Thomas K.: "A mathematical model of the finding of usability problems," Proceedings of ACM INTERCHI'93 Conference (Amsterdam, The Netherlands, 24-29 April 1993), pp. 206-213
5. Mayhew, D.J. (1999) The Usability Engineering Lifecycle: A practitioner`s handbook for user interface design. Morgan Kaufman
6. Pernice, K., Nielsen, J. (2001) Beyond Alt text: making the web easy to use for users with disabilites. `http://www.nngroup.com/reports/accessibility/`

A GOMS Model for Keyboard Navigation in Web Pages and Web Applications

Martin Schrepp[1] and Patrick Fischer[2]

[1] SAP AG, Dietmar-Hopp-Allee 16, 69190 Walldorf, Germany
martin.schrepp@sap.com
[2] University of Mannheim, Germany
fischerp@rumms.uni-mannheim.de

Abstract. An unlimited keyboard support is one of the main accessibility requirements for web pages and web applications. But it is not sufficient that the user can perform all actions on the page using the keyboard. In addition designers of web sites or web applications have to make sure that keyboard users can use their pages with acceptable performance. We present GOMS models for mouse and keyboard navigation in web pages and web applications. These models can be used to compare keyboard navigation with mouse navigation. Such a comparison allows us to decide if the amount of keyboard support for a web page or web application is sufficient or if there is an unacceptable disadvantage for keyboard users.

1 Introduction

An unlimited keyboard support is one of the main accessibility requirements for web pages and web applications. The user must be able to reach all interactive elements of a page (links, fields in online forms, etc.) by keyboard actions. In addition the user must be able to perform all possible actions on the page (follow a link, press a button, etc.) using the keyboard.

This requirement is for example covered by the WAI 1.0 [1] Guideline 9 *Design for device independence* or Section 508 [2] checkpoint §1194.21a *When software is designed to run on a system that has a keyboard, product functions shall be executable from a keyboard where the function itself or performing a function can be discerned textually.*

But it is not sufficient that all active page elements are available over keyboard. In addition a user must be able to navigate a page using the keyboard with acceptable performance.

This is especially important for web applications [3]. Vendors of business software increasingly offer parts of their software as web applications. There are two reasons for this trend. First, companies want to save costs by avoiding installations of a GUI (Graphical User Interface) on each workplace in the company. Second, mobile or distributed business scenarios require that employees can connect to the IT infrastructure of their company over the web.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 287 – 294, 2006.

The accessibility of web applications is important for the workplace integration of disabled people [4, 5]. Computerized workplaces offer handicapped persons often a good chance for a job which is not in conflict with their disability. Especially web applications have a high potential concerning workplace integration, since they allow their users to access all relevant information over the web. Handicapped users can profit from such applications since they allow them to work fully or partly from home. But this requires that the design of web applications takes the special needs of disabled users into account and allows them to work efficiently.

The design concept of *Design for All* (this concept is also often referred as *Universal Design* or *Barrier free Design*) tries to address this problem [6, 7]. The basic idea of this concept is to create user interface designs which as many people as possible can use either directly or with the help of assistive technology. Currently the *Design for All* concept seems not to be well adopted by designers. For example, a recent study [8] on web site accessibility conducted by the *British Disability Rights Commission* showed that most web sites (81%) fail to satisfy even the most basic checkpoints of the Web Content Accessibility Guidelines [1]. In addition interviews with disabled users reported in this study show that these users believe that most current web sites do not consider their specific needs.

Since users of web applications work with those applications on a daily basis they must be able to work efficiently. Thus, it is for example not acceptable that a task which can be solved using the mouse in 1 minute requires 10 minutes if the user is restricted to use the keyboard. If we fail to provide an efficient keyboard support we exclude disabled persons from working with the software, even if it is accessible accordingly to the common accessibility guidelines [1, 2].

It is important to notice here that an adequate keyboard support is not only important for disabled persons. The needs of disabled persons concerning keyboard support are similar to the needs of expert users. Such expert users often prefer to handle an application completely over the keyboard [9, 10], since working with the keyboard is much faster than working with the mouse. This need for efficiency is also covered by the common definition of usability given in the ISO-Norm 9241 (Part 11) [11] where usability is defined as the extent to which a product can be used by specified users to achieve specified goals with *effectiveness*, *efficiency* and *satisfaction* in a specified context of use. Thus, improving the accessibility of a software product by providing an efficient keyboard support will also improve the usability of the product [4].

2 GOMS Models for Navigation in Web Pages

The standard method to access interactive elements of web sites, like links or fields in forms, is provided by including these elements into the TAB-chain. Thus, all interactive elements of the page are organised into a virtual chain. The user can navigate to the next respectively previous element on this virtual chain by pressing the TAB key respectively the key combination SHIFT + TAB.

This method to make web sites accessible over keyboard is problematic if a web site contains a huge number of interactive elements, since the user has to press the TAB-key very often to reach the required element [3]. Thus, for web sites which contain a huge number of links or other interactive elements it is necessary to add accelerators, for example access keys, to provide a reasonable level of keyboard support.

2.1 Comparison of Mouse and Keyboard Navigation

We want to make sure that a user is able to navigate in a web page or web application using the keyboard with sufficient performance. A natural criterion to decide if keyboard navigation is efficient is to compare it to mouse navigation. Let k be the time required by an experienced user to finish a task using the keyboard and m be the time necessary to finish the same task using mouse and keyboard. We can define that keyboard support is sufficient if $k < c * m$, where c is a constant depending on the nature of the usage scenario. Thus, for a news page on the web obviously a higher value is acceptable for c than for a web application which is used for professional work.

Thus, we need a method to compare mouse and keyboard navigation. In principle this can be done with user tests, but this requires a huge effort. A cheap method to compare different designs concerning their efficiency is the GOMS model [12,13, 14]. The acronym GOMS stands for *Goals*, *Operators*, *Methods* and *Selection rules* which are the basic building blocks of such a model.

A GOMS model allows to predict how long an experienced user needs to perform a given task in a given interface design. There are several different GOMS models available. For an overview of these models see [13].

The central idea of the GOMS model is that the time necessary to perform a certain task is the sum of the times of the elementary actions (pressing a button, moving the mouse cursor to a certain place at the screen, clicking on a link, etc.) required to finish the task. Different users will need different times for these elementary actions. But for a comparative analysis of screen designs it is sufficient to use a set of typical or average times for the elementary actions. These average times are typically determined in laboratory experiments [13, 15].

We describe in the following two GOMS models for mouse respectively keyboard navigation in web pages or web applications.

2.2 A GOMS Model for Mouse Navigation

Assume that the task of a user is to follow a link on a given page. If the user navigates using the mouse this task can be split into three sub-tasks:

- locate the target link on the page,
- move the mouse focus to the target link,
- click on the target link.

Let t_1 be the time required to locate the target link, t_2 be the time required to move the mouse to the target link, and t_3 be the time to click on the target link. Thus, the total time to perform this task can be computed as

$$t = t_1 + t_2 + t_3 \,. \quad (1)$$

The time t_3 consists in fact of two components. First, the time the user needs to decide if the link to which the mouse points is the correct target (cognitive decision time). Second, the manual time required to click on the target link.

A good estimate for t_2 (1.1 seconds) is already published [13,15]. Estimates for t_1 and t_3 will be given later in this paper.

2.3 A GOMS Model for Keyboard Navigation

Assume again that the task of a user is to execute a link on a given page. If the user navigates using the keyboard and if we assume that the target link is the *n*'th link in the TAB-chain the task can be split into the following subtasks:

- locate the target link on the page,
- press n-times the TAB key to place the cursor focus on the link,
- follow the link by pressing the ENTER key.

There exists already a good estimation for the time necessary to press a key on the keyboard (0.2 sec) [15]. But this time can not be used directly for a GOMS model of keyboard navigation. The reason is that the cognitive task of the user requires that he or she controls the current cursor position while pressing the TAB key repeatedly. Cognitively we can assume that the process which controls the cursor position interferes with the process of pressing the TAB key. Thus, we can expect that the time required to press the TAB key is higher than the above mentioned estimate. In addition the user may loose the current cursor position completely and needs to find it again in the page.

We assume in our GOMS model for keyboard navigation that the total time to complete the task is given by:

$$t = t_1 + n\ t_4 + n\ p\ t_5 + t_6\ . \qquad (2)$$

Here t_1 is again the time required to locate the target link, t_4 is the time required to press the TAB key, p is the probability to loose the cursor focus during navigation, t_5 is the time required to locate the lost cursor focus again, and t_6 is the time required to follow the link by pressing the ENTER key.

Again we assume that t_6 consists of two components. The first component is the cognitive decision time required to decide if the link is the correct target link. The second component is the time required to press the ENTER key.

The described GOMS model concentrates on the keyboard interaction of sighted users with a web page or web application. This model is not fully adequate for the interaction of blind users with such applications. For blind users we have to consider the fact that they can scan the page only linearly over their screen reader or Braille display. Thus, the process to locate the target link and to follow this link is different for such users. GOMS models to describe the interaction of blind users with web pages are described in [16].

3 A Study to Estimate the Model Parameters

We performed a study to estimate the parameters of our GOMS models. The task of the participants was to navigate to a link on a web page using only the keyboard. The mouse was not available for the participants during the task.

3.1 Participants

10 persons participated in the study. Their age ranged from 25 to 43 years. All participants were experienced computer users. The participants were not paid for their participation in the study.

3.2 Material

To get reliable measures for the parameters it is necessary to use a wide variety of different web pages. Therefore, we picked 15 common German web pages for the study. This sample contained, for example, web shops, information pages of companies or public organizations, web pages of television channels, etc.

For each page a target link for the navigation task was identified. The number of links on these pages varied between 38 and 200. The number of TAB presses necessary to reach the target link varied between 14 and 119.

3.3 Procedure

Each participant was welcomed and placed in front of a monitor. The instruction for the participant was visible on the monitor in form of an HTML page. The participant was instructed to read this instruction carefully and to clarify open questions with the experimenter before he or she starts the experiment.

The participant was instructed that he or she has the task to navigate in web pages using only the keyboard. Navigation over the TAB key respectively the key combination SHIFT + TAB was explained in detail.

Each participant had to solve the navigation task for 10 of the selected web pages. The pages which are assigned to the participant and their sequence were determined randomly per participant.

The instruction page contains in addition direct links to the pages used in the study. These links were embedded in a short text which named the target link on the test page. The participant could start the experiment directly from the instruction page by following the corresponding links.

When the participant followed such a link, the corresponding web page was loaded and the participant could start to solve the navigation task for this page. Measurement was done in the background by a tool which was installed locally on the test computer. This tool captured all keyboard events of the participant and the corresponding system time. Thus, we are able to calculate the time necessary for each keyboard action by a difference of the captured system times between two successive keyboard actions.

After the participant has clicked on the target link in a test page, he or she was instructed to close the browser window. The participant was then automatically redirected to the instruction page. The participant could then start the next page in the sequence again over a link in the instruction page.

4 Results of the Study

Each participant had to solve the navigation task for 10 of the selected web pages. The first two pages processed by a participant were considered as trial pages, which should allow the participants to get familiar with the TAB navigation. These two pages are therefore not included in the analysis. The other 8 pages are used for an estimation of the parameters of the GOMS model.

Since each of the 10 participants solved the navigation task for 8 different web pages (the 2 trial pages are excluded) we got 80 different observations. Five of these observations must be excluded from the analysis because of technical problems with the used web pages during the experiment. Thus, the estimation of the parameters for the GOMS model is based on 75 different observations.

The participants were instructed to avoid pressing keys until they have located the target link on the page. Thus, the parameter t_1 can simply be estimated as the average time difference between the loading of the page and the first press of the TAB key. Since the manual execution time for pressing the key is still included in this estimation we additionally subtracted 0.2 seconds (usual estimation for the time required to press a key on the keyboard).

The time t_6 to execute a link over the keyboard can be estimated directly as the average time between the last press on the TAB key (which sets the focus on the target link) and the press on the ENTER key. Since the manual execution times to press a key on the keyboard and to press a mouse button are usually identical (0.2 seconds) [13, 15], we can use the estimate for t_6 also as an estimate for t_3.

We calculated per participant all differences between the captured TAB events. These differences were ordered accordingly to their size and analysed using the Nalimov test. This statistical test is designed to detect *outliers* in the data. A value is called an outlier here, if the hypothesis that it belongs to the population can be rejected with a risk of error smaller than a defined probability α (we used $\alpha = 0.1$ in our study). The detected outliers were identified as focus losses. Their average is thus the estimate for t_5 and their frequency is the estimate for p. The average of all other times is the estimate for t_4.

We got the following estimates for the parameters:

- t_1: 9.6 sec.
- t_4: 0.27 sec.
- t_5: 2.61 sec.
- t_6: 1.13 sec.
- p: 3.18%

Focus losses occur relatively seldom (in only 3.18% of the cases). But if we calculate their influence on the task completion time we see that they are responsible for 20.71% of the total time subjects needed to solve the navigation task.

In addition the most time consuming part of the navigation task is the initial location of the target. This accounts for 33.87% of the total time subjects needed to solve the navigation task.

Please note that the calculated times for the parameters of the GOMS model are valid for non-disabled highly trained computer users. Since disabled users who are not able to use a mouse will be much slower in hitting a key their disadvantage in performance compared to non-handicapped users can be in fact much higher (some authors estimate that the average time a disabled user needs to press a key is higher than 0.6 seconds [17]).

5 Conclusions

We presented GOMS models for mouse and keyboard navigation in web pages and web applications. These models can be used to compare different ways to operate a web page or web application. Especially they allow comparing keyboard navigation with mouse navigation. Such a comparison enables us to decide if the amount of keyboard support for a web page or web application is sufficient or if the disadvantage for keyboard users is unacceptable.

Assume, for example, that we have a web page with 100 links which has not implemented any shortcuts or access keys. Thus, keyboard navigation must be done solely over the TAB chain. Assume in addition that users of this web page select all links with the same probability. Accordingly to our GOMS model for mouse navigation the average time to detect and follow a link on this page using the mouse will be 11.83 seconds. According to our GOMS model for keyboard navigation the time required for this navigation task using the keyboard is 28.56 seconds. If we want to achieve that the performance of keyboard navigation should not exceed two times the performance of mouse navigation we can conclude that we must implement additional keyboard support for this page (see, for example, [3] for a description of different methods to enhance keyboard support in web pages).

References

1. Chrisholm, W., Vanderheiden, G., Jacobs, I. (1999): Web Content Accessibility Guidelines 1.0. Available online under `http://www.w3c.org/TR/WCAG10/`
2. US Department of Justice. Section 508 of the Federal Rehabilitation Act. Available online under `http://www.section508.gov/`
3. Schrepp, M., Jani, R. (2005): *Efficient keyboard support in web-pages.* In Pruski, A. & Knops, H. (Eds.): Assistive Technology: From Virtuality to Reality. IOS Press, S. 504 – 508
4. Jani, R., Schrepp, M. (2004): Influence of Accessibility Related Activities on the Usability of the Business Software. In: Miesenberger, K., Klaus, J., Zagler, W. & Burger, D. (Eds.), ICCHP – Computer Helping People with Special Needs. Volume 3118, pp. 52-59
5. Keates, S., Clarkson, P.J, Coy, J., Robinson, P. (1999): Universal Access in the workplace: A case study. Proceedings of the 5th ERCIM Workshop, Dagstuhl, Germany, pp. 73-80
6. Bühler, C., Stephanidis, C. (2004): European Co-operation Activities Promoting Design for All in information Society Technologies. In: Miesenberger, K., Klaus, J., Zagler, W. & Burger, D. (Eds.), ICCHP – Computer Helping People with Special Needs. Volume 3118, pp. 80-87
7. Stephanidis, C., Salvendy, G. (1999): Towards an information society for all: HCI challenges and R&D recommendations. International Journal of Human-Computer Interaction, 11(1), pp. 1-28
8. Disability Rights Commission (2004): The Web Access and Inclusion for Disabled People. A formal investigation conducted by the Disability Rights Commission. London: TSO
9. Cooper, A. (1995): About Face: The essentials of User Interface Design. IDG Books, Foster City, CA

10. Nielsen, J. (2000): Novice vs. Expert Users. In: Useit.com - Usable Information Technology. Available online under
`http://www.useit.com/alertbox/20000206.html`
11. ISO 9241-11 (1998): Ergonomic requirements for office work with visual display terminals - Part 11: Guidance on usability. Genf: International Organization for Standardization
12. Card, S.K., Moran, T.P., Newell, A.. (1983): The Psychology of Human-Computer Interaction. Hillsdale, NJ, Erlbaum.
13. John, B.E., Kieras, D.E. (1996): The GOMS family of user interface analysis techniques: Comparison and Contrast. ACM Transactions on Computer-Human Interaction, 3, 4, pp. 320-351
14. Kieras, D. (1997): A Guide to GOMS Model Usability Evaluation using NGOMSL. In: Helander, M., Landauer, T.K. & Pradhu, P. (Eds.), Handbook of Human-Computer Interactions. Elsevier
15. Olson, J.R. & Olson, G.M. (1990): The growth of cognitive modelling in human-computer interactions since GOMS. Human-Computer Interaction, 5, 221-265.
16. Eichstädt, H. (2005): Interaktionen blinder Nutzer bei der Bedienung linearisierter Oberflächen [Interactions of blind persons during the usage of linearized user interfaces]. In: C. Stary (Ed.), Mensch & Computer 2005: Kunst und Wissenschaft – Grenzüberschreitungen der interaktiven ART, Oldenbourg Verlag: Munich, p. 61 – 70
17. Keates, S., Clarkson, P.J. & Robinson, P. (1998): Developing a methodology for the design of accessible interfaces. Proceedings of the 4'th ERCIM Workshop, Stockholm, Sweden

People with Disabilities: Accessible Tourism
Introduction to the Special Thematic Session

Franz Pühretmair

Competence network information technology to support the integration of people with disabilities (KI-I), Hauptstrasse 119, A-4232 Hagenberg, Austria
fp@ki-i.at

Abstract. Making tourism accessible means to make tourism enjoyable for everyone, whether the tourist is a young or an old tourist, a wheelchair user, a visually or hearing impaired person, a mother with a baby carriage, a single parent family, a traveler with heavy bulky baggage or someone recovering from an accident or an illness. The accessibility itself refers to the physical accessibility of the infrastructure, facilities and services at the holiday location as well as to the accessibility and availability of information regarding on-site accessibility. Addressing accessible tourism, the tourism industry can gain an economic advantage and an added value for tourism destinations, which attract the expanded market of all tourists.

1 Introduction to the Special Thematic Session on Accessible Tourism

The philosophy of Tourism for All (Accessible Tourism) is based on the concept of Design for All. The idea behind is to design products, services and the entire environment in a way that the broadest group of users is enabled to make use of it equally. Often Design for All is misunderstood to be a "Design for people with disabilities". But the important aspect of design for all is to become part of the mainstream design practice, rather than design for a niche market of "the disabled" [1]. With respect to the tourism industry Design for All means: "Everyone, regardless of whether they have any disabilities – should be able to travel to the country, within the country and to whatever place, attraction or event they should wish to visit" [2].

But the reality shows the difference, tourists with disabilities are regularly confronted with physical, economic, social and attitudinal barriers when traveling. Physical barriers include inaccessible accommodations, restaurants, sights, visitor attractions and inadequate transport as well as inaccessible Web-based information and lack of adequate or appropriate assistance [3]. As scientific studies have shown [2] people with disabilities have the same motivation and a similar desire for traveling as everyone else. They retire from traveling only because of barriers, missing offers or missing information on accessibility features.

"Nearly half of all handicapped people would travel more frequently, if there were more barrier-free offers. And about 40 percent have already renounced a travel because adequate offers were missing"
(Max Stich, ADAC vice-president for tourism, 2003).

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 295 – 297, 2006.

The motivation for accessible tourism does not only rely on ethical and social considerations alone. The demographic change in the society will have a huge impact on the tourism industry and is therefore a significant argument to support accessible tourism. Elderly people often hit similar barriers and have requirements corresponding to people with disabilities. The market for disabled and elderly visitors has been dubbed the last major untapped market for tourism [4]. Good accessibility of tourism objects and services is not only from interest for people with disabilities, it makes tourism in general better accessible for a wide range of the population. The creation and the design of an accessible, spacious and comfortable environment should be regarded as a sign of quality, which will make journeys and holidays more convenient and pleasant not only for people with disabilities and elderly people. An accessible destination will fulfill tourists' wishes like comfort, time-economy and reduction of stress [5]. For instance a cable-car that can be entered stepless is a requirement of use for about 10% of the population. For approximately 30 – 40% a stepless access simplifies the use and for 100% of the population it is an argument of comfort. A higher number of accessible facilities and services make traveling easier and increase the comfort factor when enjoying holidays [2]. Thus raising the level of satisfaction among all tourists.

Even if the tourism industry recognizes the market potential, they will not establish full accessibility by themselves. This can only be reached if there is an according body of laws and policies. Increasingly there is a growing body of national laws and policies, like the Americans with Disability Act (ADA) [6], the Disability Discrimination Act in the UK (DDA) [7] or the eEurope action plan [8] that has caused European regulations regarding eAccessibility. With such legal and regulatory activities accessibility can be established as a civil right of all citizens.

Due to their disability, people with disabilities have special needs regarding holidays, which mainly correspond to the accessibility of transportation, accommodation, sights, infrastructure, tourism services, etc. Beside such physical accessibility requirements, the information on accessible features presented on accessible information material is another important key criterion. The strategic planning of travels does not start at the holiday location, information on accessibility influence the whole booking and decision making process [9]. Missing accessibility information or inaccessible information is often the cause why products or services do not attract tourists with disabilities and their considerable market potential remains untapped.

Offering accessibility information of a certain hotel, city or region can make tourism products more attractive for customers with disabilities and older people [9]. To establish an "accessible destination", the tourism industry has to concentrate on improvements in the following areas [2]:

- Awareness for and knowledge about problem situations that may arise on location for any customer, be they disabled or able-bodied.
- The distribution of reliable basic information on the accessibility (including its restrictions) of a destination that any guest may encounter. The spreading of such accessibility information has to be done in an accessible way like accessible tourism Web-sites.
- Ensuring guests' ability to experience the main attractions and facilities, and safeguarding the accessibility the entire tourism service chain, with a focus on facilities' functionality and ease of use.

Ideally every tourism business serious about creating an accessible environment for guests should audit their premises and present this information to their customers (e.g. on their accessible tourism website). Such an audit should be a first step to address accessible tourism. Furthermore it will help to identify areas of improvement and to set-up an action plan for accessible tourism within the business plan of the tourism service provider.

In reality, the tourism industry is still far away from offering accessible tourism to enable all tourists to enjoy equitable holidays. But there is a positive trend that accessible tourism is going to reach a far higher significance in the tourism industry. First tourism enterprises have already identified the sustainable economic chance to address accessible tourism as a criterion for quality and a competitive advantage. Investments in accessible tourism are worthwhile, because they make traveling easier and more enjoyable and lead to a higher convenience, comfort and quality for all tourists.

The Special Thematic Session "Accessible Tourism" is organized to provide a forum for the discussion of major issues related to accessible tourism for people with disabilities, to identify existing barriers as well as technologies, strategies and approaches to implement and promote accessible tourism.

References

1. Dazentas, J., Miesenberger, K., Design for All in Information Technology: A Universal Concern, Sixteenth International Conference on Database and Expert Systems Applications (DEXA 2005), IEEE Computer Society, Copenhagen, Denmark (2005)
2. Neumann, P., Reuber, P., et al., Economic Impulses of Accessible Tourism for All, Study commissioned by the Federal Ministry of Economics and Labour (BMWA), ISSN 0342-9288, Berlin, Germany (2004)
3. European Disability Forum (EDF), EDF Position Paper on Tourism: Framing the Future of European Tourism, Doc. EDF 01/13 EN, Brussels, Belgium (2001)
4. Tourism for All UK (TFA), All are Welcome!, online available on http://www.tourismforall.org.uk/index.htm (accessed on 10.04.2006)
5. European Concept for Accessibility Network (EUCAN), European Concept for Accessibility (ECA) – Technical Assitance Manual, ISBN 2-919931-24-5, Luxembourg (2003)
6. Americans with Disabilities Act (ADA), Information and Technical Assistance on the Americans with Disabilities Act (ADA) online available at http://www.usdoj.gov/crt/ada/adahom1.htm (accessed on 10.04.2006)
7. The Disability Act 2005 (DDA), online available on http://www.direct.gov.uk, (accessed on 10.04.2006)
8. eEurope 2002 – An Information Society for All, Communication on a Commission Initiative for the Special European Council of Lisbon, p. 18, March 2000
9. Pühretmair, F., It's time to make eTourism accessible, 9th International Conference on Computers Helping People with Special Needs (ICCHP 2004), ISBN 3-540-22334-7, Springer Verlag, pp. 272-279, Paris, France (2004)

Accessibility Add-on Box Enabling Barrier-Free Tourism Information Systems (TIS)

Michael Winkler and Wolfram Wöß

Institute for Applied Knowledge Processing (FAW)
Johannes Kepler University Linz, Austria
michael.winkler@liwest.at, wolfram.woess@jku.at

Abstract. Barrier-free tourism as an enormous market potential affects the whole tourism chain covering all tourism objects from arrival, accommodations, restaurants, activities on holiday resorts up to the departure process. In order to offer barrier-free tourism, two kinds of accessibility have to be considered. Tourism accessibility extends common tourism object especially with information concerning the requirements of disabled people. Web accessibility is responsible for the presentation of tourism accessibility in a transparent and accessible way. This paper introduces an *Accessibility Add-on Box* as extension for existing tourism information systems which supports both providers of tourism objects in adding extra accessibility information to tourism objects as well as tourists with disabilities in retrieving required accessibility information.

Keywords: Tourism information systems (TIS), accessibility, elderly/disabled people, barrier-free tourism, eTourism.

1 Introduction

People with disabilities often suffer from a lack of information concerning several fields. Among these fields we discover for example technical gimmicks, public traffic and in the context of this paper: tourism. Tourists with disabilities require information concerning the locality where they are willing to spend their money and they need this information before being in place. The World Wide Web as a virtual reality offers the great opportunity to discover a desired location even by multimedia from a safety and well-known home environment supported by assistive technology.

In addition the accessibility information concerning tourism objects in a tourism information system has to be presented in an accessible way. This results in a distinction between tourism accessibility and web accessibility. The requirements of both tourism and web accessibility strongly depend on the user's kind of disability.

Tourism accessibility not only covers information about the tourism object itself. Furthermore, the complete tourism chain starting with information retrieval, reservation and arrival up to events at holiday resorts has to be accessible. In this paper accessibility operationally is defined as follows: a tourism object or a web site is accessible, when people with disabilities can use them almost as efficiently as people without disabilities can do [7].

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 298 – 305, 2006.

The decision whether a tourism objectis accessible for a specific person or not has to be estimated by the person (tourist) itself and not by the provider. Tourism industry has to provide appropriate information in order to give the full decision potential to the users.

The aim of the presented accessibility box is twofold: first, it extends an existing tourism information system (TIS) with metadata information containing accessibility information concerning tourism objects, and second, it makes this information accessible within a TIS client application.

This paper is organized as follows: Section 2 gives a brief overview of related work in the fields of tourism and web accessibility. Both tourism and web accessibility in the context of the presented AccessibilityBox are introduced in Section 3. The overall concept and implementation of the AccessibilityBox add-on software enabling barrier-free tourism information systems is discussed in Section 4. Finally, Section 5 concludes the paper and provides perspectives for further research.

2 Related Work

Nowadays, an increasing number of web sites offer users the opportunity to adjust or to modify the presentation of the visited page. This covers, for example, the contrast between background and fonts, the font face, the font size, etc [6]. Other web sites prompt the user to choose between a regular edition of presentation or an accessible one - often the so-called "text only" version.

Such approaches often focus on a certain group of user with disabilities and therefore suffer from a lack of flexibility and respect against other user groups. In addition settings affecting the presentation of the complete web site often have to be decided on the very first starting page and these settings can only be modified by returning to this start page often resulting in a loss of information (the actual search result or settings concerning the main functionality of a web application). Even the loss of the contents of a shopping card can be a consequence of modified user settings if this modification results in session invalidation (a new log in). During practical work the modification of previously taken settings is frequently required, e.g. changing the contrast of a picture or the font size.

Contrarily to the approach of letting the user customize parts of the design patterns other web sites provide accessibility information in tourism context for a selected group of users (for example wheelchair users). Several online hotel guides provide the tourism objects with quality standards like "seniors", "children" or "wheelchair users" [11]. Often the checklists resulting in such a quality standard are invisible for the user and therefore not transparent. Web accessibility and usability are no main objectives of such tourism information systems and therefore these systems often are difficult to use and navigate for people with disabilities [4]. Furthermore tourism accessibility information often is cut down to attributes like "near downtown" [4].

DEHOGA - barrier free tourism and restaurants - developed quality standards for accommodations and restaurants in Germany for different user groups [3]. Additionally, checklists contain at least 10 issues for each user group, which support tourism offer providers.

Key concepts of the AVANTI project [8] in tourism context focus on adaptability and adaptivity both at the user interface and the content level. In order to provide the information required by users with disabilities and elderly users, the AVANTI browser retrieves and integrates tourism data from heterogeneous multimedia and none-multimedia databases. Beyond the opportunity to fade in tourism information relevant for users with disabilities, no means for an additional specification of the required tourism accessibility information are offered.

The user groups supported by DEHOGA [3] mainly differ from the potential user categories developed by ICF (International Classification of Functioning, Disability and Health) [12] and both user sets do not come in line with the potential user groups developed by W3C (World Wide Web Consortium). The WCAG1.0 (Web Content Accessibility Guidelines 1.0) [10] developed by the WAI (Web Access Initiative) proposes a user matrix with relevant checkpoints for different user groups and scenarios how people with disabilities use the web [9].

All approaches of formally evaluating a document in order to guaranty accessibility suffer from the fact that accessibility is not specified in the formal document structure but in the user’s experience while browsing a web site or using the functionality of a web application [7]. It is possible to produce Web resources that conform to WCAG 1.0's Priority 1 checkpoints and comply with Section 508's standards for Web accessibility but still do not make sense to people with disabilities [7].

The Inclusive Learning Exchange (TILE) [2] demonstrates the potential benefits of personalization for people with disabilities. Multimedia data stored in a repository are presented to the users according to their preferences and transformations (e. g. font face, size, colour and a toolbar). ATutor [1] is a working Learning Content Management System (LCMS) sharing data with TILE and keeping accessibility and adaptivity in mind. Both systems are supported by an authoring tool to manage the multimedia information and e.g. to identify alternative resources for a certain presentation.

MultiReader [5] encounters the "lost in hyperspace" phenomenon in focusing on navigation and orientation in large multimedia documents (e. g. the WWW). To achieve these goals, MultiReader implements a non-visual, a visual-only and a multimodal user interface (covering all kinds of print disabilities) with possibilities to individualize these interfaces.

A very first initial stage of solutions for an AccessibilityBox where the user selects a fixed and predefined profile (for example "deaf user") and therefore only tourism objects accessible for deaf users are included in the result sets can not fulfill the demands concerning flexibility, self-determination and individuality. Even single user groups are characterized by different requirements for tourism accessibility as well as web accessibility. Furthermore, there a two other important aspects: web accessibility guidelines have to be implemented independently from a selected profile and people with multiple disabilities are unable to choose a fixed profile.

These considerations and difficulties in existing web applications leaded to the general idea of the presented accessibility box.

3 Accessibility for Tourism Information Systems

In the context of web applications supporting barrier free tourism two types of accessibility have to be distinguished: tourism accessibility and web accessibility which both affect certain kinds of users with disabilities and their different requirements.

Based on the quality standards of barrier free tourism and restaurants (DEHOGA) [3] the present application distinguishes three user categories: visually impaired and blind users, hearing impaired and deaf users and finally going impaired people and wheelchair users. This distinction is of theoretically nature only, because a certain user is free to check any accessibility attribute of any user group. This fact is important to meet the requirements of free self-determination and to support users with multiple disabilities.

3.1 Tourism Accessibility

An AccessibilityBox should implement at least tourism accessibility attributes of the modules mentioned in the tourism chain above. With these few modules on the one hand a user can simply plan accessible holidays, on the other hand a tourism provider has the means to offer accessible holyday packages. Furthermore the model can easily be extended for example by city areas like downtown, airport, rail station etc. covering certain accommodations, restaurants and of course accessibility attributes of city areas like public wheelchair toilets or guiding systems for blind people.

3.2 Examples of Tourism Object Accessibility Attributes

The attributes of tourism objects are no disjunctive subsets of the complete set of accessibility attributes. For example, a restaurant excludes attributes from accommodations and adds new attributes to that subset.

Accessibility attributes of accommodations and restaurants are for example: a high contrast and bright illumination, an accessible online bill of fare, an inductive system, the opportunity of table reservation by mail or fax, handicapped parking lots, desks navigable with wheelchairs, etc.

3.3 Presentation of Tourism Objects

The description of a tourism object covers the following information types:

- A video file, an audio description, an image and a describing text in any language supported (file size and extension restricted to default settings).
- A detailed geographic location of the tourism objects, e.g. Germany/Hesse/ Frankfurt/Neu-Isenburg - Frankfurt/Rhein-Main, whereas the appendix Frankfurt/Rain/ Main is the region.
- Links (and mail-to links) to the provider's homepage, to the city's homepage and to the tourism object's homepage. Furthermore, a link to Google, which specifies a corresponding query string including the selected language.

3.4 Web Accessibility

The WCAG1.0 [10] published by the W3C consists of 14 guidelines and 65 checkpoints. The importance of a checkpoint is expressed as its priority (level A to triple A) [10]. The US Section 508 is mainly based on the WCAG1.0 checkpoints with priority A [6,7].

These static checkpoints are relevant for any kind of user for any web site or application and thus are followed throughout the AccessibilityBox as far as appropriate. The set of users in the present application includes visitors of the tourism web site as well as tourism providers using the extranet that also can be a person with disabilities. In addition to these static guidelines the user has the opportunity to dynamically modify the contents and the layout of the presentation by using the toolbar of the AccessibilityBox as follows:

- In the user interface access keys are defined for push buttons. The access keys can be turned off to avoid conflicts with (1) opening the menu bar of the user's favorite browser and (2) shortcuts of the user's assistive device.
- Users can change the language of the application at any time and in any use case. To demonstrate multilingual use of the AccessibilityBox the present implementation supports English, German and traditional Chinese. Today's screenreaders used by blind people are able to automatically switch the language of the included speech synthesizer if the user changes the language of the AccessibilityBox.
- Images can easily be zoomed and faded away. Additionally, users have the possibility to invert the screen in order to modify the contrast of the site.
- To support visual impaired people, the used font family and the font size can be changed at any time of application runtime.

4 Accessibility Add-on Box Implementation

The overall idea and intention of the present AccessibilityBox software is the presentation of accessibility information in tourism context for all user groups and additionally offering means for adapting the presentation of this accessibility information whenever required or desired by a so called toolbar.

The introduced AccessibilityBox acts as extension of an existing tourism information system: the implementation, data structures and used techniques are aiming at a minimal adaptation as well as maintenance effort for the existing TIS. For example, the initialization of additional data corresponding with the tourism objects of the TIS, which is required by the AccessibilityBox, is done automatically during runtime.
Typical use cases of the AccessibilityBox are as follows:

- Provider design the accessible presentation of tourism objects
- Users register to the AccessibilityBox and manage their profile
- Registered users as well as anonymous users browse and query the TIS database.

Tourism providers are allowed to add accessibility information (accessibility attributes) to their tourism objects in the extranet of the web application while users themselves establish their personal profile by checking which accessibility attributes they require.

Both, anonymous users without specified tourism accessibility attributes as well as registered users with such specifications in their profile are supported in browsing or querying the database required for tourism information. The presentation of the query results should be as flexible as possible. The AccessibilityBox therefore allows a user to minimize the font size on a certain page displaying a single query result and at the same time to zoom the provided image and to invert the screen. Even at this point the language of the entire web application including its user interface, any text and ALT-text attributes as well as the multimedia presentations can be changed easily.

Furthermore and very important for users with multiple disabilities, detailed information concerning a tourism object, its accessibility and its provider can be faded in easily.

4.1 The ExtraNet - Adding Accessibility

After a successful extranet login in the first step a tourism provider will find a table showing a list of tourism objects that are offered by the tourism company. Each row of this table corresponds to one tourism object showing the status of the multimedia files already uploaded.

This table of tourism objects covers the entire information according to the language selected in the toolbar. Changing the language on this site not only has an impact on the language of the user interface but also on the data displayed in the table: the table dynamically provides the status of the multimedia files according to the selected language in the toolbar. Features like selecting the amount of rows to be displayed on a single page and restricting or expanding the categories of tourism objects to be displayed are available on this page. Dynamically generated links allow the navigation through the result pages.

By selecting a certain tourism object the tourism provider enters the upload and working area where the presentation of a tourism object can be designed. Above this upload and working area the presentation area shows exactly the same presentation of the selected tourism object that a visitor of the web site will see. Changes made in the upload and working area immediately influence the presentation area. The tourism provider has the opportunity to upload video and audio files with the accepted extension and limited file size. A link text for an uploaded file is mandatory. Analogous, images can be uploaded; the attributes of the images like width, height and ALT can be modified at any time including a textual description for a tourism object.

Changing the language on this extranet site affects both the user interface and the tourism object's presentation: the provider designs the presentation of the tourism object according to the selected language.

Furthermore the accessibility attributes of the selected tourism object are checked or unchecked which is directly visible in the user interface.

4.2 Using the AccessibilityBox

Visitors of the tourism web site can take advantage of the functionality in the role of anonymous and registered users. The functionality of both anonymous and registered users is exactly the same with the exception that only tourism objects corresponding to the registered user’s settings are included in any result set. A registered user specifies exactly the same set of tourism attributes and toolbar settings as mentioned in Section 3 and these settings can be modified by the user at any time in a so called “MyBox”.

A user will be recognized in the case of a further page visit, if cookies are allowed on his/her browser and therefore tourism objects will be selected according to the user settings. For any modification of tourism accessibility settings a login by the user is mandatory. Independent of the toolbar settings stored in the user’s profile the current screen settings will be recovered at a further visit if cookies are activated on the user’s device.

The tourism objects contained in the database can be browsed by category, e.g. accommodation or restaurant (an implicit abstract level covering all categories is also available). Furthermore, a user can search the whole database for keywords. Then the query includes all attributes of the tourism objects as listed in Section 3.3. In order to improve the usability and convenience a user can specify the number of items to be displayed on a single result page showing a list of tourism objects as well as the category of tourism objects to be displayed.

4.3 Accessibility Add-on Box Implementation

Implementation principles and data structures aim at flexibility and a minimum of maintenance effort in introducing new tourism categories, introducing further languages and adding new attributes to existing categories. Accessibility attributes do not act as attributes of a tourism object but are stored in a separate table and therefore can be maintained regardless any queries.

Table 1. Tourism accessibility attributes

Tourism_object	Accessibility_attribute	Value
43001	Illumination	0
43001	Online bill of fare	1

The association of accessibility attributes and tourism category is enclosed in the object structure and occurs at runtime whereat accessibility attributes are stored in a dynamic data structure and thus they do not affect any method interface.

This information mapped to XML, stored in Java Beans and formatted with XSLT enables a dynamic generation of forms to retrieve tourism object's accessibility attributes on the one hand and tourism accessibility attributes specified by the user on the other hand.

To exclude a tourism object queries search for a single accessibility attribute specified in the user's profile but not checked in the tourism object's specification. Therefore tourism objects, which are already inserted in the TIS database but not yet adapted by the tourism provider in the extranet immediately, appear in the user's result set. This flexibility is achieved because there simply is no accessibility attribute to be compared with accessibility attributes specified in the user's profile and consequently the result set is not cut down.

5 Conclusions and Future Work

This paper introduces the intelligent Accessibility Add-on software, which extends existing tourism Information systems: tourism providers add accessibility attributes to their tourism objects and design the accessible presentation of these objects (multimedia). Users of the AccessibilityBox retrieve accessibility information required in an accessible way and organize the overall presentation according to their individual needs in a so-called toolbar.

Future work has to focus on testing the acceptance of the AccessibilityBox software among users with different disabilities. In addition, a universal model of accessibility attributes outside the tourism context based on the resource description framework (RDF) standard will be developed. Such a universal model could significantly extend the application scenarios. For example, a technical gimmicks provider who uses an Internet portal for offering products obtains the means to add corresponding accessibility information (e. g., accessible instruction manuals, usability information, etc.) to the items offered.

References

1. Adaptive Technology Resource Centre (ATRC): ATutor, http://atutor.ca/ (2006)
2. Canarie Inc. Ottawa, Adaptive Technology Resource Centre (ATRC): The Inclusive Learning Exchange (TILE), http://www.barrierfree.ca/tile/ (2006)
3. Deutscher Hotel- und Gaststättenverband e.V (DEHOGA Bundesverband): DEHOGA - Barrierefreies Gastgewerbe – barrierefreier Tourismus, http://www.dehoga-berlin.de/home/barrierefreiheit_1039_1009.html (2005)
4. Expedia Inc Washington, Expedia.com GmbH München: Expedia.de, http://www.expedia.de/ (2006)
5. Information Society Technology (IST): MultiReader, http://www.multireader.org/ (2006)
6. IT Accessibility & Workforce Division (ITAW): Section 508: The Road to Accessibility, http://www.section508.gov/ (2005)
7. Slatin John M., Rush Sharron: Maximum Accessibility: Making Your Web Site More Usable for Everyone, Addison Wesley (2002)
8. The Avanti Consortium: AVANTI, http://www.ifac.cnr.it/avanti/contents/home.htm (2006)
9. The World Wide Web Consortium (W3C): How People with Disabilities Use the Web, W3C Working Draft, December 2004, http://www.w3.org/WAI/EO/Drafts/PWD-Use-Web/ (2005)
10. The World Wide Web Consortium (W3C): Web Content Accessibility Guidelines 1.0, W3C Recommendation 5-May-1999, http://www.w3.org/TR/WAI-WEBCONTENT/ (2005)
11. Tiscover AG Innsbruck, FAW Hagenberg: Tiscover, http://www.tiscover.at/ (2006)
12. World Health Organization (WHO): International Classification of Functioning (ICF), Disability and Health, October 2001, http://www3.who.int/icf/icftemplate.cfm (2005)

Development of a Mobile Tourist Information System for People with Functional Limitations: User Behaviour Concept and Specification of Content Requirements

Sascha M. Sommer[1], Marion Wiethoff[2], Sari Valjakka[3], Dionisis Kehagias[4], and Dimitrios Tzovaras[4]

[1] Institute for Occupational Physiology at the University of Dortmund, Germany
sommer@ifado.de
[2] Technical University of Delft, Netherlands
m.wiethoff@tbm.tudelft.nl
[3] Research and Development Centre for Welfare and Health, Finland
Sari.Valjakka@stakes.fi
[4] Centre for Research and Technology Hellas / Informatics and Telematics Institute, Greece
{Dionisis.Kehagias, Dimitrios.Tzovaras}@iti.gr

Abstract. The paper describes the specification of content requirements for a new mobile tourist information service for people with functional limitations. The theoretical background for the specification of content requirements is based on action and activity theory. The framework enables the division of complex activities into smaller functional units in order to analyse the information needs of user groups with different types of functional limitations. The approach provides results in the form of information elements and attributes which can be easily translated into a machine-readable language.

1 Introduction

The transdisciplinary project ASK-IT is developing an ambient intelligence system that provides assistive tourism and leisure-related information to people with functional limitations.[1] Users of the system can ask, at any time of day, for information about transportation, hotels, venues or events with a mobile phone or a PDA-like device. Service users will receive tourism and leisure-related accessibility information tailored to their personal user profile. This paper describes the route from an activity-centred specification of service content requirements to the translation of the identified requirements into a machine-readable format.

People with functional limitations face constraints and barriers in all stages of touristic or leisure-related trips [1]. Useful and timely information about mobility barriers and suitable touristic offerings to overcome them could reduce these problems. The

[1] ASK-IT: Ambient intelligence system of agents for knowledge based and integrated services for mobility impaired users. Integrated project co-funded by the Information Society Technologies programme of the European Commission. Project number IST 511298; duration 2004 – 2008).

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 306 – 313, 2006.

information needs for every goal-directed action depend generally on the complex interaction between, on the one hand, the individual (physical abilities, psychophysiological capacities, cognitive resources etc.) and, on the other hand, relevant factors of the environment (objects in a scene, available tools, implicit and explicit context rules etc.). Riva suggested accordingly to focus on relevant user activities when analysing requirements for ambient intelligence environments [2]. The psychological frameworks action and activity theory are approaches to conceptualize goal-directed human behaviour. Action theory enables the division of complex actions into smaller behavioural units [3]. Activity theory stresses, moreover, the social context of human behaviour [4].

2 Theoretical Framework for the Specification of Requirements

The three building blocks of the theoretical background for the specification of content requirements for the new service are the definition of user groups together with the elaboration and implementation of relevant action and activity theory principles.

The user groups are classified on the basis of *functional* limitations. For instance, whether a person is unable to use the upper limb is very important for the requirements analysis, but whether the functional limitation is caused by an injury or a rheumatic disposition is, for this analysis, of minor importance. ICF codes that take into consideration the interaction between health conditions and contextual factors hence provide an adequate basis for the definition of user groups that has been proven to be appropriate in previous projects [5, 6]. User groups were defined accordingly in two stages: first a main group classification, and second a nested sub group classification of different levels of severity. In this way, diversification of user groups is guaranteed, while the option to cluster sub-user groups in a sensible way for reasons of parsimony is still available. The one exception is the wheelchair user group which is classified as a separate main user group, because their functional requirements differ considerably from other users with lower limb limitations. The resulting user group classification has the following main groups: (1) lower limb impairment, (2) wheelchair users, (3) upper limb impairment, (4) upper body impairment, (5) physiological impairment, (6) psychological impairment, (7) cognitive impairment, (8) vision impairment, (9) hearing impairment, (10) communication production/receiving impairment.

Action theory provides a general framework for human behaviour in working environments, which can be applied to any type of goal-directed human behaviour [2]. The flow of actions and operations to achieve a goal is organized by higher levels of action regulation. The action process as a whole is hence organized by hierachical-sequential action patterns. Subordinate actions and operations are controlled by personal goals and sub-goals. Fig.1 (next page) shows, as an example for a hierarchical-sequential action process, the flow of actions and operations of a person arriving by car at a hotel.

Decomposing complex goal-directed actions in this manner enables, with sufficient detail, to identify specific support needs of users with different types of functional limitations. E.g. a person with lower limb impairments would need mainly information about the way from the car park to the hotel, while a wheelchair driver would need, in addition, information about the type of doors to the hotel lobby.

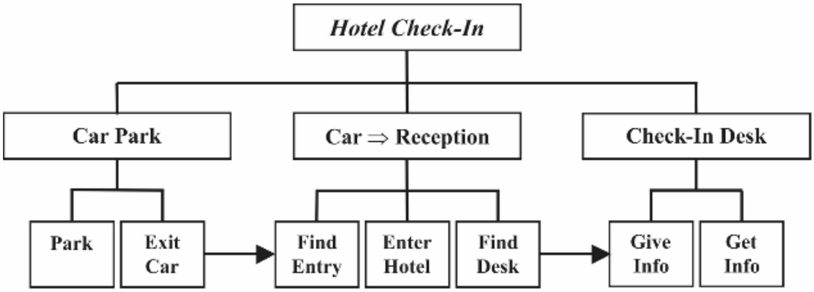


Fig. 1. Example of a hierarchical-sequential action process: A person traveling by car arrives at a hotel

A shortcoming of action theory is the neglect of the social context and the artefacts in different environments. Activity theory allows describing the so-called "minimal meaningful context" into which an action is embedded. The most important elements of the context are the object(ive) of the activity, the subjects involved, the available tools and the division of labour, the set of rules under which the action takes place, and the community in which the subject takes part (Figure 2; [7,3]).

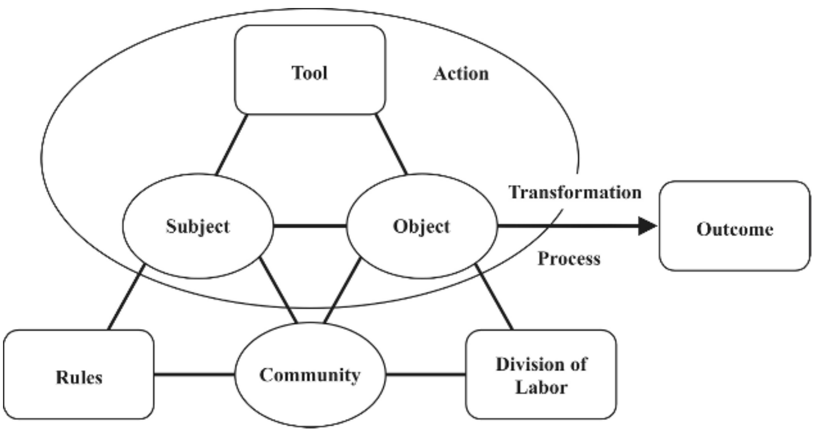


Fig. 2. Representation of an activity as described by [3]

The *object* of the activity is the element that participants manipulate and transform, and which turns into an outcome (goal). E.g., when the object(ive) is to check-in at a hotel the desired outcome is to acquire the room key. Important issue in activity theory is that the transformation process itself, and hence the role of each element in the minimal meaningful context, should be analysed. The context-based representation of the interaction between users and the environment, with a succession of activities and a hierarchy of actions and operations, allows the structured definition of content requirements as needed for the new mobile tourist information service.

3 Implementation of the Theoretical Framework

The aim of the new mobile tourist information service is to provide assistance at two levels: During action/activity planning (e.g. which hotel to choose), and during the execution of plans (e.g. operations to get from the car park to the entrance). Trip planning comprises activities of getting motivated, finding information and making pre-arrangements [see e.g. 9]. When executing the plans during the trip the following general activities need to be considered: Getting to points of interests, getting into the facility, using the facility, using the services and getting out in an emergency. *General actions* are common for all facilities, whereas *specific actions* are optionally applied depending on the type of facility and the offered service. Information requirements common for all facilities include, among others, *general actions* like arriving by public transport or car, parking, walking/wheeling along outdoor routes to the entrance and moving from one room to another. In addition, general tourist information on various points of interests and services is needed. Information on available services essential to people with functional limitations (first aid, medical services, assistance services, ICT services etc.) is also required.

Specific actions such as sleeping (accommodation), eating (restaurant, café), shopping (shop), taking part in an event (spectator's stand), visiting an exhibition and making sports (swimming pool, beach, walking trails) etc. are related to further specific accessibility information requirements.

Relevant requirements were specified in a matrix for each user group. Table 1 presents an extract of the matrix for user group 2 (wheelchair users).

Table 1. An extract of the tourism and leisure matrix for wheelchair users

1. Activity	*2. Action*	*3. Attribute*	*4. ValueType*	*5. ValueLimit*
Facilities: Getting in	**Accessing entrance**	**Stairs**	**Boolean yes/no**	
		Thresholds	**Boolean yes/no**	
		Height of thresholds	**mm integer number**	**maximum 20 mm**
		Width of doors	**mm integer number**	**minimum 850 mm**
		Ramp	**Boolean yes/no**	**yes, if stairs**

The matrix structure was developed in three steps:

1. Identification of activities and distinction of actions: Activities relevant for tourism and leisure in both stages were identified, and broken down into actions in the sense of action theory (columns 1 and 2). For instance, if an objective of a person travelling by car is to check-in at a hotel, the actions of parking the car, getting from the car park to the entrance, accessing the entrance, and checking-in at the reception need to be performed.

2. Definition of a set of attributes: Activity theories' minimal meaningful context conception was applied as a framework for the definition of information requirements by user groups. The attributes describe in a structured way the environmental factors, which make / do not make accessible operations possible. To each action a set of user group specific attributes can be mapped (column 3). Environmental attributes cover characteristics of the built environment (e.g. type of door), tools (e.g. availability of assistive devices) and community related aspects (e.g. if help is available by staff).
3. Specification of details of attributes: Attribute types, attribute values and value limits based on accessibility guidelines were specified for each user group (columns 4 and 5). Attributes of Boolean type (e.g. a yes/no answer to the question if a ramp is available) or numeric attributes (e.g. a gradient measurement) were favored. User group specific limit values reflect average requirements of a specific user group [e.g. 8]. In the development of the new mobile tourist information service, these limit values can be utilised for the creation of user group specific profiles.

4 Content Modeling Procedure

The goal of the content modeling procedure is to provide a formal description of user information needs in a computer understandable and interoperable format. The application of the theoretical framework resulted in content requirements presented in table format (see Table 1). These information requirements are the basis for the content modeling procedure. The outcome of the modeling procedure is a set of computer-interpretable models that represent the user information needs. These models describe the pieces of information that should be exchanged between the user and different data sources or heterogeneous applications. By imposing a set of constraints on data, a common delivery format is dictated. Thus, when a user with a functional limitation requests a new service, the common data format, which acts as an information filtering facility, guarantees that the user gets access only to valid data values. XML was chosen for representing models, because it is by far the most supported content representation language today. An XML-schema is a definition of the structure of an XML document. Given an XML document and a schema, a schema processor can check for validity, i.e. that the document conforms to the schema requirements.

Table 2. Division of information elements into attributes and their description

Information Element	Attributes	Explanation
Bus station	- Station Id	A unique identifier
	- Accessible Place	A list of accessible places attached to the bus station
	- Entrance	The point of entrance to the bus station
	- Schedule	The schedule of itineraries of the bus station
	- Address	The postal address
	- Working Hours	The opening hours of the bus station
	- Buses	A list of elements of type *Bus*

The procedure for moving from the content requirement matrices to the XML schemes involves the transformation of the matrices into a *tree* structure, consistent with the notation of an XML schema. Each concept related to a specific user information need is encoded as an information element composed of several attributes, related to values of information that the user desires to know in order to be able to participate in tourism-related activities. In Table 2 (previous page), an example of one *information element* and its *attributes* is shown.

The next step is to create the corresponding XML-Schema document. The latter is actually a representation of a collection of elements as the one described in Table 2. A graphical representation of an arbitrary information element comprised of three attributes, is illustrated in Figure 3. This tree-like graphical representation is provided by the XMLSpy authoring tool, which supports automatic generation of XML code by knowledge engineers, without requiring a deep knowledge of XML. The corresponding XML-Schema document that describes the element of Fig. 3 is given in the code segment of Fig. 4. The author creates the tree, which is illustrated in Figure 3 using a set of an appropriate graphical user interface, while the authoring tool automatically generates the code shown in Figure 4.

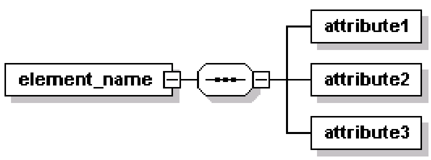


Fig. 3. A XMLSpy-generated graphical representation of an element with three nested elements as attributes

```
<xs:element name="element_name">
 <xs:complexType>
  <xs:sequence>
   <xs:element name="attribute1"/>
   <xs:element name="attribute2"/>
   <xs:element name="attribute3"/>
  </xs:sequence>
 </xs:complexType>
</xs:element>
```

Fig. 4. An element with attributes in XML-Schema

In the example shown in Table 4 each element is specified by the xs:element identifier. Since the element 'element_name' consists of three elements (i.e. more than one other element), it is a complex data type. This is specified by the identifier *xs:complexType*. The *xs:sequence* literal represents an ordered collection of elements that describe the list of three attributes that constitute the 'element_name' in Fig. 4. Primitive XML data types can also be encoded in any schema in order to describe simple data values of the various structural elements. For example, in order to represent temperature in Celsius degrees, the *xs:decimal* type, which is suitable for representing decimal numbers, would be used. The application of the transformation of the content requirements into the XML tree-like structures will eventually result in tree representations as the one shown in Figure 5. It illustrates how a specific element (i.e. a bus station, see Table 2), which is composed of a list of other nested attributes (entrance, schedule etc.), may form the basis for the definition of a new custom complex data type.

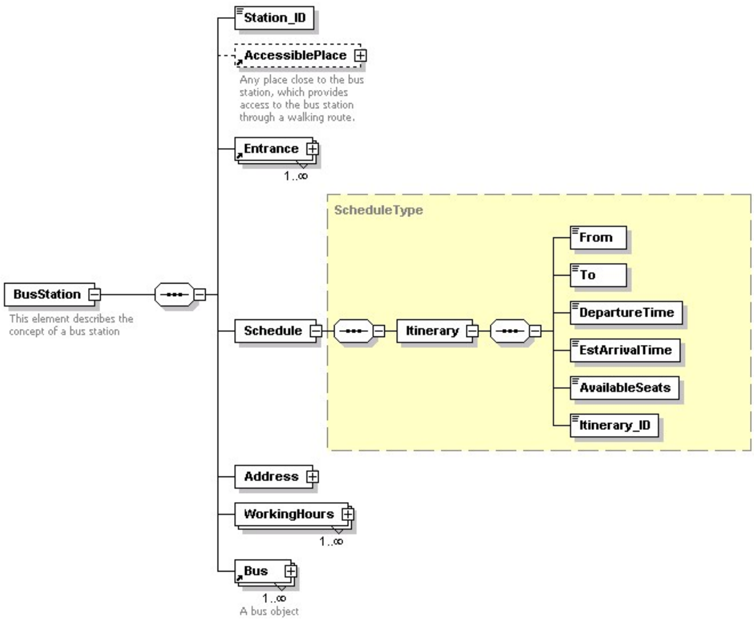


Fig. 5. XMLSpy-generated graphical representation of the BusStation element

5 Discussion and Conclusions

The above described content definition and modeling procedure has been successfully applied in the framework of the ASK-IT project. The matrix structure based on action and activity theory principles facilitated the systematic and extensive content requirements specification. Comprehensive content requirement matrices were produced. The tables include user group specific attributes for all identified actions and activities. A difficulty was the wide variety of relevant tourism and leisure-related activities, and hence the huge quantity of information needs related to their performance by people with various types of functional limitations. One of the main challenges was thus to keep the list of information needs and attributes manageable. For this purpose, essential actions common for all facilities and specific optional actions depending on type of facility and offered services were distinguished. The extensive lists of content requirements were, furthermore, evaluated and prioritised by representatives of the different user groups. Only those attributes with high priority for the user group in question were translated into XML schemes.

Morganti and Riva emphasised recently that the focus of Ambient Intelligence for rehabilitation should be the support of the users' activities and interactions with the environment [12, p. 285]. The integration of action and activity theory principles has indeed proven to be a suitable theoretical framework for the specification of content requirements for an ambient intelligence-based tourist information service for users with functional limitations. The content requirements matrices provide, for each user group, a structured representation of information elements in the form of classes with attributes and limit values. This approach facilitates the subsequent creation of XML schemes, because the input for the content modeling procedure is immediately available in a format that can be converted into a machine-readable language without difficulties.

References

1. United Nations: Barrier-Free Tourism for People with Disabilities in the Asian and Pacific Region. United Nations Publications, New York (2003)
2. Riva, R.: The Psychology of Ambient Intelligence: Activity, Situation and Presence. In: Riva, G. , Vatalaro, F., Davide, F., Alcaniz, M. (eds.): Ambient Intelligence, 17 – 33, IOS Press Amsterdam (2005)
3. Frese, M., Zapf, D.: Action as the Core of Work Psychology: A German Approach. In: Dunnetee M. D. et al. (eds.): Handbook of Industrial and Organizational Psychology, Vol. 4, 271 – 340, Consulting Psychologists Press Palo Alto (1994)
4. Kuutti, K.: Activity theory as a potential framework for human computer interaction research. In: Nielsen, J. (eds.): Usability Engineering, Academic Press London (1993)
5. World Health Organization: International Classification of Functioning, Disability and Health (2001))
6. TELSCAN project: Inventory of ATT System Requirements for Elderly and Disabled Drivers and Travellers, Deliverable 3.1 (1997)
7. Engeström, Y.: Learning by expanding. Helsinki: Orienta-Konsultit (1987)
8. Aragall, F.: European Concept for Accessibility (ECA). Technical Assistance Manual. (2003)
9. Psymarnou M., Ekberg J., Valjakka S., Hanninen S., Hofmann J., Bekiaris A., Maglavera S.: Preliminary user requirements and priorities. Sport4All Deliverable D1.1. Provision of Telematic Services to the Disabled for the Diffusion of Information on Athletic Events. (2003)
10. Morganti, F., Riva, R.: Ambient Intelligence for Rehabilitation. In: Riva, G. , Vatalaro, F., Davide, F., Alcaniz, M. (eds.): Ambient Intelligence, 283 – 295, IOS Press Amsterdam (2005)

Objective Accessibility Assessment of Public Infrastructures

Christophe Ponsard[1] and Vincent Snoeck[2]

[1] CETIC Research Center, Charleroi, Belgium
cp@cetic.be
[2] GAMAH Association for Better Accessibility, Namur, Belgium
vincent.snoeck@gamah.be

Abstract. More than 30% of the population are experiencing daily problems to access public infrastructures. Despite this, there is still no effective and widely adopted method to measure accessibility, inform users and incite managers to improve their infrastructure.

The Passe-Partout Index was designed which such goals in mind. Starting from a thorough accessibility requirements analysis accounting for obstacles related to various impairments, a complete set of measurable criteria and evaluation rules were developed. Those enable an precise, relevant, informative and objective assessment of the accessibility of public infrastructures.

1 Introduction

Everybody has or will experience, at some point of his life, a state of reduced mobility: a broken leg, travelling with a big case or wandering with a baby car. Suddenly a door, stairs or some device becomes an insurmountable obstacle. Unfortunately some of us experience this every day because of a physical, sensorial or cognitive impairment: people in wheelchair or with crutches, blind people, deaf people, etc. It is estimated that about 30% of the population are experiencing daily problems related to accessibility.

While everybody agrees on the fact that accessibility has to be improved, it is however difficult to get a precise picture of accessibility due to a lack of well-defined evaluation methodology. Over time, the widely used ICTA logo [8] (a white wheelchair on a blue background) has lost most of his meaning due to poorly defined attribution criteria, his restricted scope (only considering physical disability) and the lack of enforcement. As detailed in section 2, such errors have been repeated time and again by various logos and accessibility labels.

Since 2002, a new methodology is being developed and experimented in Wallonia. It is led by GAMAH, an 25 year old association fighting for a more accessible world. This method stands out by a precise and global approach, assessing the accessibility for all kinds of mobility impairments and of public infrastructures (administration, shopping, tourism...). It aims at a double objective: (1) inform users about accessibility and (2) urge infrastructure managers for improvements. The approach was designed using a very careful requirements engineering approach where obstacles were identified from potential disabilities, requirements

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 314–321, 2006.

were grouped, prioritized and structured in a coherent way, resulting in the attribution of a value between 0 and 9 for 6 categories of people with special needs [9]. It is called the "Passe-Partout Index" (PPI) and is depicted as a key showing the accessibility score (see figure 1). The process followed to define PPI and the key on how to use it are detailed in section 3.

Fig. 1. Scope of the FAUST toolbox

The precise underlying accessibility model was used to effectively deploy the method: first the reference document was written, then paper forms, training material and more recently computer support were developed for systematically taking notes during visit, help in the evaluation process and to write reports. This is explained in section 4.

2 Review of Accessibility Assessment Methods

This section reviews a number of accessibility assessment methods developed over time in a number of countries. Those are depicted in figure 2: the wheelchair logo from ICTA (international) [8], Tourism and Handicap (France) [3], Acces City (Belgium) [2], Kéroul (Canada) [6] and Accesmetrie (France) [1]. Our purpose is not to be exhaustive but to point out a number of problems using the following evaluation grid: approach of disability, application domain, precision, enforcement and communication tools.

Fig. 2. Logos of a number of accessibility methods

Approach of Disability. In the past, the focus was mainly on the physical handicap. So the wheelchair logo from ICTA [8] became international symbol of access. As other disabilities - related to hearing, visual and cognitive impairments - have specific accessibility requirements, the need to make then more visible appeared over time. The majority of assessment methods now support various impairments and use specific logos for them.

Application Domain. A number of methods target tourism, for example "Tourism and Handicap" [3]. A recent European study [13] also focus on tourism. Those are oriented to the evaluation of specific infrastructures such as hotels, restaurants, sites (museum, monuments,...) or entertainment (theatre, sport,...). While a good driver to raise awareness, they leave aside a number of infrastructures related to administration, health/social care, teaching, shopping and transportation.

Precision and Enforcement Level. Those topics are related and discussed together.

The *ICTA logo* is not enforced at all. Although some guidelines are available [4], nobody checks them. This is a reason why it has lost most of his meaning except maybe for toilets, lifts and parking.

Labels such as "Tourism and Handicap" provide guidelines and rely on self-evaluation based on forms. They are generally quite imprecise and easy to get. Control is also rarely seriously enforced. As a result, only few managers feel an incentive to improve the accessibility of their infrastructure.

More precise methods rely on a well defined set of criteria, generally of very similar nature, although with some variants (e.g. for the exact measure required for a door width or a slope). Some methods define explicit levels (e.g. Acces City). Other methods quantify over some scale which can then be interpreted (e.g. Accessmetrie). Figure 3 illustrates this.

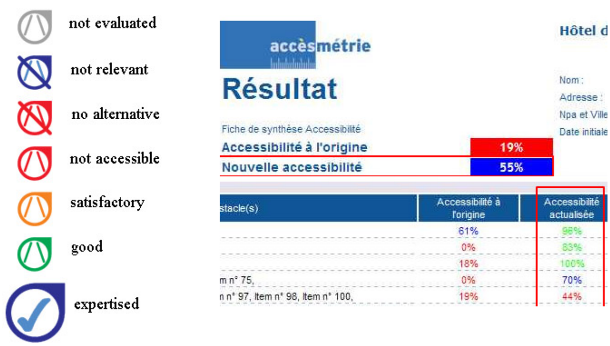


Fig. 3. Representations for Acces City (left) and Accesmetrie (right)

Precise methods require expertise to use them. They are generally applied or coordinated by a central organization which also ensures a good enforcement level by independent evaluation and diffusion. Previously free of charge and managed by association of disabled people (e.g. for Acces City), the current move is towards a commercial service (e.g. Accessmetrie) where infrastructure managers have to pay for the evaluation and advices. This is not a barrier as managers are now quite aware of accessibility. Having to pay a fee also results in a greater involvement. Recognition by authorities can improve visibility and even make evaluation a mandatory step for getting a building or renovation authorisation.

Communication Tools. Information is a key issue. Most methods use logos which are applied on the infrastructure and diffused in various publications such as tourism guides. For the general public, the representation should be both simple and intuitive to decode, by relying on standard logos (wheelchair, eye, hear...) and colours conventions (see figure 3).

A dedicated website is also very frequent: this ensures a broad, searchable and up-to-date diffusion of the information. To be coherent, the design of the website should also cope with computer accessibility issues especially for visually impaired people.

As conclusion of this short review, a good method should target all kind of mobility impairment for all kinds of public infrastructures. It should rely on a very precise set of pertinent requirements meeting the needs of a wide scope of impairments. This precision should not translate into complex communication: appropriate representation to inform must be available. Finally, the method should also be enforced by a dedicated organization recognized by local authorities.

3 Design of the Passe-Partout Index

A requirements engineering approach was used to design a method meeting the above conditions [11]. A complete, pertinence and measurable set of requirements were discovered and structured through the explicit modelling of goals, users abilities/limitations and obstacles to overcome. Such an approach was already applied with success to derive personal requirements for impaired people [10].

3.1 Goal Analysis

The first step of the method was to define a goal model. Such as model capture both the functional and non-functional goals that people visiting a public infrastructure want to satisfy. Functional goals are systematically refined and structured using various operators such as temporal patterns (such as $Park \Rightarrow Enter \Rightarrow Circulate \Rightarrow Use$) and case-based decomposition ($Vertical\ Circulation$ versus $Horizontal\ Circulation$). Non-functional goals covers requirements such as security, non discrimination, autonomy and service availability.

3.2 Obstacle Analysis

Obstacles to the above goals were identified and structured using knowledge provided by experts and reasoning based on physical, sensorial and cognitive impairments[12]. For example, for a wheelchair, limitation on the width can result in problems to pass through a door, the nature of the chair forbids the use of stairs and the strength required to operate the chair will limit the ability to climb slopes.

Those obstacles have then to be removed using various strategies. A new requirement may be introduced to avoid them. For example the requirement on a minimal door width to allow wheelchair to pass through it. Another strategy is to propose an alternative way to satisfy the goal: the staircase obstacle can be removed if an alternative accessible way to reach other levels is provided using a ramp or a lift for example. Of course, those could introduce new obstacles, for example if the alternative way is through a remote secondary door which would discriminate the person. In the end, this results in a rich set of accessibility requirements.

3.3 Category Definition

The next step was to define categories reflecting consistent set of requirements with which most people can identify. Six categories were defined (see figure 4). The first three categories are devoted to physical impairments (1) wheelchair alone, (2) with help and (3) walking people. The two next categories are devoted to sensory impairments: (4) visual and (5) hearing. The last category is devoted to cognitive impairments (6).

Those categories are associated with a target group and have been validated by relevant associations active in the related domain. In the grouping process some assimilation occurred for less represented disabilities. For example, little people have same height requirements as wheelchairs and have to identify with category (2). Generally speaking when using the PPI, the first step is to know in which category is relevant (possibly more than one). For example, some deaf people (category 5) also have a limited level of written language and will take benefits of simplified explanations provided for cognitive disabilities (category 6).

Category	Requirements	Who can identify ?	
1 – Wheelchair alone	Heigth – Width – Security	Person in wheelchair alone (eg. paraplegic) Little people	
2 – Assisted wheelchair	As for wheelchair alone but relaxed. Physical aid	People in wheelchair with reduced autonomy (eg. tetraplegic) requiring assistance.	
3 – Walking problem	Limited autonomy, Rest – Security	People with crutches, stability trouble, People easily tired (asthma, hearth patient) Elderly people	
4 – Visually impaired	Acceptation of assistive dog - Security – Adapted Signalling and Equipment.	Blind people People with low sight	
5 – Hearing impaired	Adapted communication devices - Trained staff – Access to sign language interprets	Deaf People Hard-of-hearing people	
6 – Cognitive impaired	Simplified information Help	Mentally handicaped people Deaf People (especially using sign language) Child, old people easily lost, foreigner	

Fig. 4. Categories

3.4 Evaluation

The final step is to be able to assign a score to measure the overall satisfaction of the accessibility requirements in each category. The choice was to give scores between 0 (not accessible at all) and 9 (excellent accessibility). For computing scores, each requirement has to be measurable. Such measures can be of different nature: observation of the existence of a required element (e.g. display of the opening hours), measure of a physical element (e.g. the width of a door), or observation of behaviours (e.g. are dogs accepted ?).

To infer the score for a category, rather than averaging the satisfaction on the relevant requirements, prioritization was considered. This allows the reader to gain direct information from the score: the 0 to 9 scale can be seen as a chain and each score is associated with a specific subset of requirements. Rules for partial evaluation were considered [7]. The score given is directly linked to the first broken requirement. The prioritization choice does generally follow the overall logic of the visit but can also account for some critical issues. Those are the most important for the infrastructure manager to fix and a low score will generally be effective for that. For example, for blind people, the acceptance of the assistant dog will come first. For wheelchairs, the priority is based on *Enter* ⇒ *Circulate* ⇒ *Use* ⇒ *Park*. *Enter* was judged critical while *Parking* a less important issue. Based on this, an example of evaluation is given in figure 5.

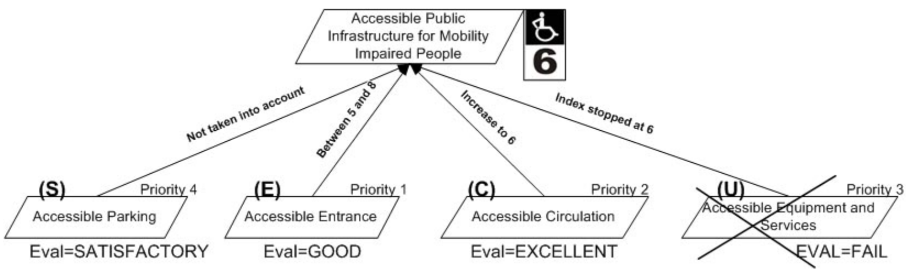


Fig. 5. Eval

One might argue that interpreting such numbers is quite complex and requires a deep understanding of the PPI, especially for the common user. This is true but the PPI can also be used informally for drawing qualitative information is enough (2=poor, 7=quite good...). A easier usage is to check for the requirements one needs and to find the "minimal" PPI in the relevant category: it is then elementary to check if an infrastructure is accessible or not !

4 Deployment

The resulting PPI is a method rich of more than 200 criteria specifying what to measure and how to compute the accessibility. The PPI is not just a set of guidelines: it is enforced by a whole organization lead by GAMAH who aims at producing and maintaining an updated view of accessibility in Wallonia. The process includes several steps: infrastructure survey, analysis, validation, diffusion and evolution. The precision of the PPI enabled a rich supporting framework to be developed.

Paper Forms. A first work was to design evaluation forms used when visiting public infrastructures. The information was easily identified using the reference document [9]. The evaluations forms follow the overall spatial structure of the index but merge elements of all categories. To draw the attention on the corresponding category, a specific icon is drawn. About 30 forms are available for evaluating parking, doors, stairs, lifts, rooms, showers, kitchen, etc. Those are assembled and filled during the visit order. Before leaving, a check is made to make sure the whole infrastructure was covered and no important information was missed.

Electronic Forms. Paper forms are quite heavy to fill in and provide little guidance. To improve this and to setup the first brick of a computer-based processing, electronic forms were developed. Used on a Tablet PC, such forms allow the surveyor to easily organize and navigate through his work and, within each form, to focus his attention on the crucial information to acquire. Based on the information collected, a first rough estimation of the accessibility can be produced immediately.

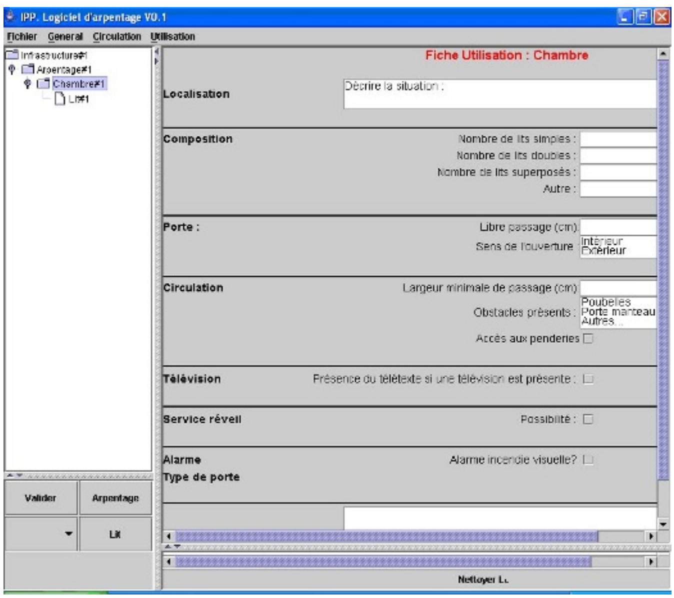


Fig. 6. Forms for surveying infrastructures

The software was developed in Java and was deployed and tested on tablet PC. It relies on a MySQL relational database both for local persistency and official repository. A snapshot of the user interface is shown in figure 6.

Diffusion. The central repository contains now more than 500 entries all over Wallonia and is accessible on the Internet through the following dedicated website: http://www.ipp-online.org with full reports and search capabilities [5].

5 Perspective and Conclusion

The Passe-Partout Index is a method for assessing the accessibility of public infrastructures based on objective, precise, complete and measurable criteria. Currently deployed in Wallonia, those unique qualities allow the method to be applied at a larger scale and to become an effective instrument to information about and to control the accessibility of public infrastructures. The PPI has also the ambition to become a framework which can help together several assessment methods to converge for the interest of all impaired users.

The IPP is still under active development. Future work is to provide more support, especially for helping the analysts to derive the scores and to write reports both for the end users and the infrastructure managers. Other developments include the use of WiFi digital camera with Tablet PC to improve picture management and to use geolocation information in search engines and for mobile deployment.

Acknowledgement

Huge credits to Christian Baeke, the father of PPI. This work was financially supported by the Walloon Ministry of Social Affairs and Health. We thanks several Walloon associations which contributed to the development and validation of the Passe-Partout Index, among them: ASPH, ACIH, CBPAM, Ligue Braille, FFSB, AFrAHM.

References

1. Accesmetrie, *Website*, `http://www.accesmetrie.com`.
2. Acces City, *Website*, `http://www.anlh.be/accescity`.
3. Secrétariat d'état au tourisme, *Label tourisme et handicap*, 2001.
4. French National Comitee for Disabled People, *Criteria and conditions for the attribution of the international accessibility symbol*, 1983.
5. GAMAH, *IPP Online*, `http://www.ipp-online.org`.
6. Kéroul, *Website*, `http://www.keroul.qc.ca`.
7. E. Letier and A. van Lamsweerde, *Reasoning about partial goal satisfaction for requirements and design engineering*, Proc. FSE 2004 - 12th Int. Symp. on the Foundation of Software Engineering, ACM Press, November 2004.
8. International Commission on Technology and Accessibility, *International symbol of access*, `http://www.ictaglobal.org/isa.html`.
9. C. Ponsard, *Specification of the Passe-Partout Index (in French)*, GAMAH, 2004.
10. A. G. Sutcliffe, S. Fickas, and M. M. Sohlberg, *Personal and contextual requirements engineering.*, Proc. 13th IEEE Int. Conf. on Req. Eng., August 2005.
11. A. van Lamsweerde, *Goal-oriented requirements engineering: A guided tour*, Invited minitutorial, Proc. RE'01, August 2001.
12. A. van Lamsweerde and E. Letier, *Handling obstacles in goal-oriented requirements engineering*, IEEE Trans. on Soft. Eng. **26** (2000), no. 10.
13. J. Westcott, *Improving information of accessible tourism for disabled people*, European Commission, June 2004.

BAIM–Information for People with Reduced Mobility in the Field of Public Transport

Christian Bühler[1], Helmut Heck[1], Dirk Sischka[1], and Josef Becker[2]

[1] Forschungsinstitut Technologie-Behindertenhilfe (FTB)
der Evangelischen Stiftung Volmarstein, Grundschötteler Str. 40,
58300 Wetter / Ruhr, Germany
baim@ftb-net.de
http://www.ftb-net.de
[2] Rhein-Main-Verkehrsverbund GmbH (RMV), Postfach 14 27,
65704 Hofheim a. Ts., Germany
baim@rmv.de

Abstract. The German project BAIM aims at supporting or enabling the active and independent participation of people with reduced mobility in public transportation by the provision of usergroup-oriented accessible information services on barrier-free travelling opportunities. Existing regional and nation-wide information services are enhanced with static and dynamic information on the accessibility of public transportation vehicles, buildings and other facilities which form part of a barrier-free travelling chain. Information and services will be accessible before and during the travel.

1 Introduction

About 8.1% of the population in Germany, i.e. about 6.7 Mill. people, are officially registered as “severely impaired”; in the European Union about 25.3% of the population are “severely hampered” (9.3%) or “hampered to some extent” (16.0%) in daily activities by any physical or mental health problem, illness or disability [1]. The number of people with reduced mobility is estimated at about 20–25% of the population (or even more, depending on the definition of mobility reduction) and is increasing due to the current demographical development in society [2]. Those people who are regarded as mobility impaired in the context of public transportation systems comprise various user groups with quite differing abilities and requirements: physically disabled people, persons with sensorial impairments, people with the reduced ability to produce or to understand spoken or written language, mentally disordered people; but also people with heavy luggage, baby carriage, or bicycle may face a “temporary” mobility handicap in public transportation.

Although many improvements towards a barrier-free public transportation system have been achieved during the past 20 years in Germany, e.g. barrier-free design of railway stations and bus stations, better accessible busses and trains, visual and auditive information systems in the vehicles, and guidance systems at station buildings, there are still many facilities that have barriers for one or the other traveler group. It

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 322 – 328, 2006.

will take some more decades to achieve a (nearly) complete accessible public transportation system.

However, people with reduced mobility could more easily avoid or overcome barriers when travelling, e.g. from one city to another, if they had information about actual barriers on their potential travelling paths and if they were supported to find "barrier-free travelling chains". A travelling chain means the whole sequence of buildings, vehicles, as well as logistic and informational infrastructure that the traveler visits and uses for a journey.

It is the goal of the German project BAIM to support or to enable the active and independent participation of mobility impaired people in public transportation by the provision of usergroup-oriented, accessible information services on barrier-free travelling opportunities.

2 State of the Art in R&D and Application

Many information systems on public transportation exist, differing with respect to region, content and way of information provision. Some systems are able to provide most up-to-date information via WAP or SMS. Typically the provided content of information aims at the not disabled traveler and is sometimes augmented with selected information for people with restricted mobility. Other information systems focus on disabled tourists and their needs, but are very selective with respect to the area or way of travelling.

There are a number of related international research projects: The European project ASK-IT [3] aims at the provision of an ambient intelligent multi-agent system that supports mobility impaired people with information on transportation, tourism, and social activities. The project Mobile Tourist Guide aims at the European touristic market and provides the user with a broad spectrum of personalised information and services. A portable terminal connects the user with the GSM/GPRS and Internet based system. In the project Capitals ITTS, an integrated travel and trourism service is being developed by inter-connecting the existing mobile service platforms of five European cities. TramMate is a mobile information service that provides a route planner for public transportation, delivers also real-time information, and supports the time planning of travelling. In the Vienna-SPIRIT [4] project an intermodal interoperable travelling information system has been developed that combines dynamic and location dependent information on various public transportation systems. The Open-SPIRIT [5] project developes corresponding electronic information devices based on smart phones.

3 Research and Methodological Approach

Within BAIM an information system is being developed that supplies people with reduced mobility with detailed information on the accessibility of public transportation facilities, including vehicles, buildings, available technical facilities, and assistive features like guidance systems for blind people. This information will be provided at the time of planning a journey at home (via an Internet based information platform or a telephone based speech information system) as well as during the journey (via an Internet/WAP service or a SMS service) so that the disabled traveler is informed in "real-time" and can react on unforeseen obstacles, e.g. a not working elevator at a station where he intended to change trains. Furthermore the BAIM information system will search for barrier-free transportation chains from any given starting point to any goal of the journey – according to the user's personal requirements profile.

BAIM is a kind of model project, that means, a concept and methodology is being developed that shall be applicable in all regions of Germany (and similar regions in Europe). The implementation and practical test will be in two selected regions which have different characteristics of their public transportation system:

One region covers the federal capital Berlin and the surrounding federal state Brandenburg. The corresponding Verkehrsverbund (public transport association) Berlin-Brandenburg has one centre (i.e. the city of Berlin) and a surrounding area with a significant lower density of traffic infrastructure ("mono-centre"). The other region is covered by Rhein-Main-Verkehrsverbund with several cities (Frankfurt, Offenbach, Wiesbaden, Darmstadt, …) and an extended rural surrounding ("poly-centres").

BAIM faces a number of research challenges:

A special challenge is the consideration of distinct requirements by different groups of travelers with mobility restrictions. Which information is essential or just useful for which user group? – Of course, the single traveler is interested in transportation chains that have no barriers for himself even if there are barriers for any other travelers with mobility restrictions. – How can user profiles be structured? – How can features of transportation facilities be classified so that a computer-based route planning will be supported?

How should information and services be presented to the users, taking into account that

1. the single user shall not get lost in a mass of information which is not relevant for him,
2. information and services shall be accessible not only at stationary PCs but also at mobile devices (PDA, mobile phone),
3. mobility impaired persons may be handicapped in the operation of PCs, PDAs and mobile phones and need alternative access devices or ways of information presentation?

How can the mass of data (with respect to accessibility or barriers) of thousands of buildings and types of transportation means be acquired and maintained efficiently?

Looking at the different sub-groups of people with reduced mobility: which kind of mobile devices do they already use and to which extent are the devices available in the sub-groups?

The first part of the solution is the fact that the BAIM information system and service will be built upon existing information platforms for public transportation. The HAFAS information server will be the backbone of the BAIM information services.

The second part of the solution is the expertise within the project consortium: two public transport associations and their related information technology service companies, the provider of the German railway information system, a company specialized in speech dialog systems, and a research institute for assistive technology. This expertise is complemented by users and user representatives with extensive experience and knowledge about the needs of disabled people concerning public transportation systems.

The third part of the solution is the methodological approach of several cycles of design, implementation and test phases during the project runtime. This will give the opportunity of gradually building up the BAIM system with early feedback from user tests.

4 Findings

A first user requirements analysis has been conducted with representatives of various groups of disabled people, all of them well familiar with problems of disabled people in public transportation, and with rehabilitation experts.

The following classification of user groups has been defined in order to elaborate the users' requirements and their priorities:

- people not being able to walk,
- walking-impaired people,
- people with impairments of the trunk and/or of their upper extremities,
- extraordinary small or tall people,
- blind people,
- visually impaired people,
- deaf people,
- hearing impaired people,
- speech impaired people,
- cognitive/mentally impaired people,
- people with organic mal-function,
- people with allergies.

As a result, the need to improve the information provision for people unable to walk, blind people and deaf people was rated with highest priority. Information for these user groups will be implemented first.

People with the unability to walk need very detailed information to enable them to judge whether they can use public transport or not.

The most important attributes should be the search criteria for barrier-free travelling chains, i. e. the conditions for connections. These attributes are gaps and steps, ramps, lifts as well as width of doors and opening hours of station entrances and other passages, manouvering space, areas for wheelchairs, multi-functional areas in vehicles. Additional information is necessary, e.g. about: toilets, counters and service facilities in stations and stops, accessible control and communication systems.

Blind people should get all information, which is given visually, in an alternative accessible way, e.g. acoustically. Interesting are, for example, dialog systems applying natural language technologies.

For deaf people all information that is normally given acoustically should be presented visually. Most important during the journey is information about delays and changes of trains and platforms, which is normally given via loudspeakers. It has to be discussed whether some information must be presented in sign language.

In order to cope with the mass of data that describe accessibility / barrier aspects of transportation facilities for the different groups of people with reduced mobility on the one hand and the requirement for usability of the information itself (especially when presented on mobile devices) on the other hand, there is a clear need for user profiling. Although very individual profiles would in principle yield the optimal search results of barrier-free travelling chains, it will not be practicable for most users to fill in long templates of parameter lists to generate their personal profile.

A possible solution is the use of user profile categories. A first distinction can be made depending on the situation when the information is used: before travelling in order to prepare the journey, or while travelling in order to receive most up-to-date information on changes or just for orientation. A second distinction for categorisation is based on the user group or the type of disability, respectively. Each group has its own information needs, of course with overlaps. While the information attributes for each disability group can be defined, including a ranking of their relevance, it is difficult, if not impossible, to further categorize certain value ranges of attributes (Table 1).

Table 1. Examples of attributes, value ranges and presettings

Attribute	Choice / value range	Presetting for wheelchair user	Presetting for walking impaired user
Weight (incl. wheelchair)	≤ 120kg ≤ 250kg ≤ 300kg ≤ 350kg	≤ 250kg	–
I need a width for passages of...	> 70cm ≥ 80cm ≥ 90cm	≥ 90cm	≥ 80cm
I can use a ramp with a gradient of...	No ≤ 6% ≤12% ≤ 20%	≤ 6%	≤ 6%
Possible average speed ...	0,5 m/s 1,0 m/s, 1,5 m/s 2,0 m/s	1 m/s	1 m/s

The reason to attempt this was the observation that an obstacle, e.g. a step, a gap or a ramp, could be a real insuperable barrier for one wheelchair user while another wheelchair user overcomes it with a small extra effort. The consequence of, for example, declaring a step, a gap or a steep ramp as a barrier for wheelchair users is that the corresponding travelling chain is not regarded as “barrier-free” and therefore would not be taken into account by the automatic route finding/planning service. In principle it is possible to define more or less arbitrary attribute value ranges for “no / low / medium / high / very high” steps etc.; but it is not really possible to reasonably cluster the ranges of many attributes to few categories. – So, the simple classification “with barrier / barrier-free” according to applicable standards can be realized in the first implementation phase at any case. Categorisation for user profiling will be prepared by the definition of attributes and preliminary value ranges (Table 1), but it needs further reasoning and trials with users.

5 Future Plans

It is intended that a first subset of functionality and information will be implemented and tested in summer 2006.

First services will base on target scheduling times and other static data for regional transportation. This data can be taken from planning tools, data bases for stopping points and for vehicle management and will be stored in the "infopool" system of plan data (Fig. 1).

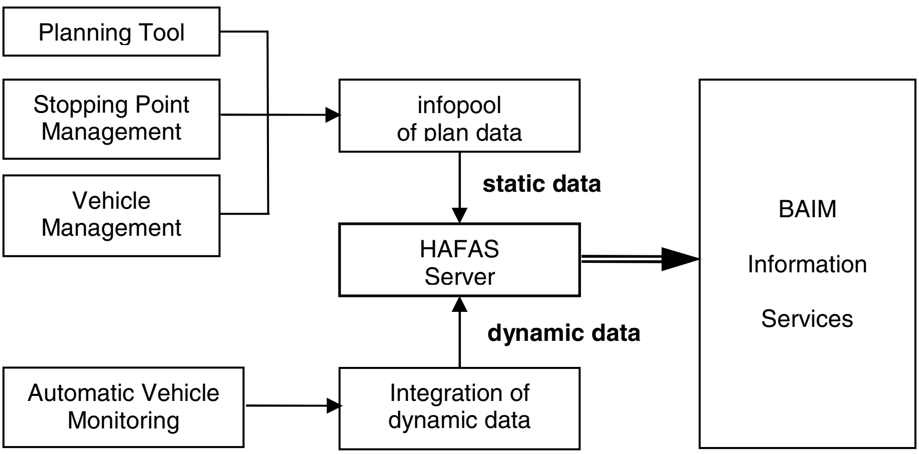


Fig. 1. Data flow in the BAIM system

At a later stage of the project, in summer 2007, dynamic information will be supplied in the information service. An important pre-requisit is the Automatic Vehicle Monitoring (AVM) which has already been introduced in Germany. It provides detailed real-time data of all vehicles. The data will be integrated by the HAFAS real-time server.

The third phase of the project will focus on the planning of barrier-free supra-regional and nation-wide journeys.

At the end of the project the findings and experiences will be wrapped up in such a way that they can be transferred to other public transport associations.

Acknowledgement

The project BAIM is co-funded by the German Federal Ministry for Education and Research (reference: 19 P 5025F) during its runtime, i.e. September 2005 to April 2008.

Further project partners, besides RMV and FTB, are Verkehrsverbund Berlin-Brandenburg (vbb), Rhein-Main-Verkehrsverbund Servicegesellschaft mbH (rms), IVU Traffic Technologies AG, HaCon Ingenieurgesellschaft mbH, and SemanticEdge GmbH.

References

1. ECHP UDB, Eurostat, 06/2003 http://epp.eurostat.cec.eu.int/portal/page?_pageid=1996,45323734&_dad=portal&_schema=PORTAL&screen=welcomeref&close=/health/public/care_his&language=de&product=EU_MASTER_health&root=EU_MASTER_health&scrollto=0
2. Verband Deutscher Verkehrsunternehmen (VDV), VDV-Förderkreis: Barrierefreier ÖPNV in Deutschland – Barrier-Free Public Transportation in Germany; Köln, 2003 (bi-lingual)
3. ASK-IT www.ask-it.org
4. Vienna-SPIRIT http://www.salzburgresearch.at/research/projects_detail_e.php?proj=72
5. Open-SPIRIT http://www.salzburgresearch.at/research/projects_detail_e.php?proj=84
6. IbGM Institut für barrierefreie Gestaltung und Mobilität: Survey on information for people with reduced mobility in the field of public transport. Study for EMTA (European Metropolitan Transport Authorities); Mainz, 2003 (Download: www.emta.com)
7. Verband Deutscher Verkehrsunternehmen (VDV), VDV-Förderkreis: Telematik im ÖPNV in Deutschland – Telematics in Public Transport in Germany; Köln, 2001 (bi-lingual)

Accessible Information Space to Promote Accessible Tourism

Franz Pühretmair[1] and Klaus Miesenberger[2]

[1] Competence network information technology to support the integration of people with disabilities (KI-I), Hauptstrasse 119, A-4232 Hagenberg, Austria
fp@ki-i.at
[2] Institute Integrated Studies, Johannes Kepler University of Linz
Altenbergerstraße 69, A-4040 Linz, Austria
klaus.miesenberger@jku.at

Abstract. Currently about 10-20% of the population are affected by a disability from temporary or permanent nature. For these people accessibility is a requirement and a necessity. The implementation of accessibility concerns all areas of he society, including tourism objects and tourism services. Besides being a social demand, accessibility has an economic dimension and is an indicator for quality, a trademark and a competitive advantage. So far, the tourism industry has hardly recognized the economic dimension of accessibility. Often the lack of not supporting and promoting accessibility is a combination of missing knowledge about requirements and needs of people with disabilities and the missing of standardized methods to evaluate accessibility, categorize and map it to standardized accessibility labels.

1 Motivation

People with disabilities have a right to enjoy travel and leisure experiences [1]. Everyone, regardless of whether they have any disability or not, should be able to travel to the country, within the country and to place, attraction or event they wish to visit [2]. As for a huge part of the population, traveling and leisure activities are important aspects in the quality of life for people with disabilities, too. However the reality shows the differences, the travel experiences of disabled tourists are often affected by transportation constraints, inaccessible accommodation and tourism sites, and inadequate customer services including inaccessible Web-sites [3] and incomplete information materials.

Studies have pointed out, that people with disabilities have a clear wish for traveling [2]. As shown in Germany [2], about half of all people with disabilities would travel more frequently if more accessible offers were available. More than a third of all people with disabilities have already decided not to travel because of missing offers, inaccessible environments or missing accessibility information. But people with disabilities form a huge market potential, because normally they are not single travelers, they are used to travel together with friends, family members or travel companions like care assistance. Furthermore people with disabilities share their market segment with the continuously

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 329 – 336, 2006.

growing group of elderly people, which often have similar requirements and needs. Additionally, the travel wants of people with disabilities are largely independent of the season and are characterized by a relatively high travel spending.

2 The Three Stages of a Travel

The needs of people with disabilities are rather divers, but the main interests always focus on the accessibility of some tourism facilities. In Germany a study has been carried out [2] to identify the facilities that have highest priority for people with disabilities. The following figure shows the results of the study, where the accessibility of the accommodation has the highest ranking followed by moving around at the holiday location and arrival and departure which implies the accessibility of the means of transport.

Importance ↑

82% Accommodation
76% Moving around at the holiday location
74% Arrival / Departure
71% Excursions
71% Organisation of the trip
62% Cultural activities
61% Arrival / familiarisation
58% Services at the holiday location
52% Medical care
51% Gastronomy
37% Shopping
19% Sport

Fig. 1. Importance ranking [2]

When analyzing a journey, three stages can be identified [4,5]:

- Travel planning and decision making, which takes place weeks or months before traveling
- Arrival and departure with different means of transport
- The stay including all activities during the stay

2.1 Travel Planning and Decision Making

Because of their special requirements, the travel planning of people with disabilities normally is characterized by a more detailed and more complex information enquiry than known from people without disabilities. People with disabilities require information with respect to their individual special needs, which are often significantly different from the information usually requested from people without disabilities [6]. Frequently the availability or the unavailability of this information influences the strategic planning of a trip and therefore the whole booking and decision making process.

The non-existence of such information results often in the fact that a city or a region is less attractive for people with disabilities. In Germany nearly 75% of the destinations offer some information on accessible facilities. But the information is often to general like the wheelchair pictogram to mark accessible facilities. Furthermore the information is often included in less available and less accessible media like general tourism brochures, but not included in the Web-portals of tourist boards' [2]. Detailed information that focuses the special interest of disabled tourists is rarely available. Additionally most state-of-the art tourism information pages and destination management systems are not accessible [7] because Web-designers did not follow the rules and guidelines for designing accessible Web-pages [8].

Therefore three general facts can be identified, that influence the travel planning and decision making process:

- Missing information about accessibility and accessible facilities
- If there is some information available, the information is mostly to inaccurate or to general
- The tourism industry still lacks in understanding the needs of these customer groups, that leads to incomplete or inaccurate information, offers and services

2.2 Arrival and Departure

As shown in Fig. 1 the accessibility conditions of the means of transport (Arrival / Departure) are a very important factor for people with disabilities. Often means of transport are characterized by low accessibility, especially public transport like trains or busses and taxis [2]. Consequently, it is often not possible to find a complete accessible transport chain. Sometimes this is the cause, why potential tourists do not book a journey or are not able to visit a certain region.

2.3 The Stay

The tree stages of a travel are hierarchical whereas the arrival and departure is split into two parts before and after the stay. The crucial part of the stay is often the information phase during planning and decision making because of a big lack in getting information about accessibility features as described in the section "Travel Planning and Decision Making".

During the stay, the interest of the traveler is focused on criterions like:

- Accommodation
- Moving around at the holiday location
- Excursions
- Cultural activities
- Services at the holiday location
- Medical care
- Gastronomy
- Shopping, sports, etc.

In relation to Fig. 1 it can be seen that three of the four most important criterions have a definite coherence with the stay. The accommodation has been identified as the most important facility during the stay. People with disabilities have a clear preference for hotels [2], because hotels often offer more accessibility features than smaller guest houses, flats, camping sites or youth hostels. Nevertheless normally there is a high level of dissatisfaction and uncertainty with respect to the accommodation because of many barriers and a lack in information about accessibility already in the planning phase.

Often the barriers have a physical nature. Furthermore, there is a frequent lack of adequate or appropriate assistance, unavailability of appropriate leisure and event packages and the absence, insufficient availability or inaccessibility of information. While many countries have national regulations and standards addressing access to public buildings and public transport there is a lack of standards for the private sector addressing hotels, gastronomy, attractions and local transport. Additionally, many barriers faced by tourists with disabilities are due to missing skills or the missing knowledge of the tourism industries employees. A better sensitization and special trainings would prevent from hitting most of the barriers.

3 Accessible Information Presentation

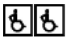
Typically each hotel, each region or destination has glossy brochures, which are available in travel agencies or must be ordered first. Such brochures are not accessible for all tourists and often do not reflect the latest state-of-the-art. Nowadays most of the tourist information is already present on Web-pages. The Internet opens the access to information which is otherwise hard to obtain and build the basis to present the most up-to-date information to all customers.

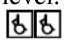
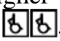
In principle there are two different approaches for a successful support of accessible information presentation in the Internet [6]. The first approach is to extend the traditional homepage of a tourism object or a destination management system with accessibility information added in an additional section or respectively on additional Web-pages. It offers the user a compact overview on accessibility information, but lacks because the accessibility information is separated from the service or product description. The second approach is to integrate the accessibility information into the original product or service description. In this way accessibility information is omnipresent and therefore easy to retrieve. For both approaches it is essential that the whole information representation is implemented on accessible Web-pages with an optimized usability [9] to make it utilizable for tourists with various prerequisites, abilities, preferences and needs.

A crucial factor is how to present accessibility information or accessibility features on Web-pages or information material. Currently it is popular to use the wheelchair pictogram to mark accessible facilities, but this is to general and imprecise. On the other hand an in-depth qualitative assessment and a descriptive listing of technical factors as it is done in the internet information service "you-too" (www.you-too.net) of the EC project Barrier Info System (BIS) is to complex. Via standardized questionnaires the accessibility features of hotels, restaurants, sights, etc. have been collected, have been stored in a database system and can be enquired. The listing of all these technical factors is to complex and therefore not utilizable for tourists when they do there enquiry.

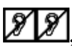
Till now there is no standardized accessibility label to determine how disability symbols should or can be used [1]. The Upper Austrian government together with the KI-I and official representatives of different disabilities is currently developing a project to establish accessibility labels. The idea is to implement accessibility labels according to accessibility conformance levels of the World Wide Web Consortiums (W3C) Web Content Accessibility Guidelines (WCAG) [8] for different kinds of disability. E.g.: for mobility impaired users, the wheelchair symbol could be used as an accessibility label.

Accordingly it is planned to define three different accessibility levels for mobility impaired people:

-  fully accessible with power wheelchair, without assistance
-  fully accessible with a wheelchair, power wheelchair needs assistance
-  basically accessible, for full use assistance is partly required

Each accessibility level summarizes a number of checkpoints [10]. Each checkpoint is assigned to an accessibility level. The compliance to the lower levels is the basis for the higher levels. To reach  it is necessary to be -conform and to fulfill all level -checkpoints. According to WCAG and to the BIS-Project, the in-depth assessment will be done with a standardized questionnaire checking all accessibility checkpoints.

Referring to the example of mobility impaired users similar accessibility labels [10] will be defined for other kinds of disability like:

- , ,  - to indicate the accessibility for blind and visually impaired tourists
- , ,  - to indicate the accessibility for deaf tourists and tourists with hearing disability
- etc.

Such accessibility labels would establish a clear standard as well as unequivocal rules to use disability labels to demonstrate the level of accessibility for all types of disabilities. Thus gives people with disabilities a clear impression of the accessibility of a hotel, a sight, a restaurant, an event location or whatever.

The challenge is the process to acquire the data and to keep the data up-to-date. Fig. 2 gives an overview on such an evaluation process. The process itself can be split into five steps of evaluation:

- Specification
- Object Selection
- Pre-Evaluation
- Detail Evaluation
- Re-Evaluation

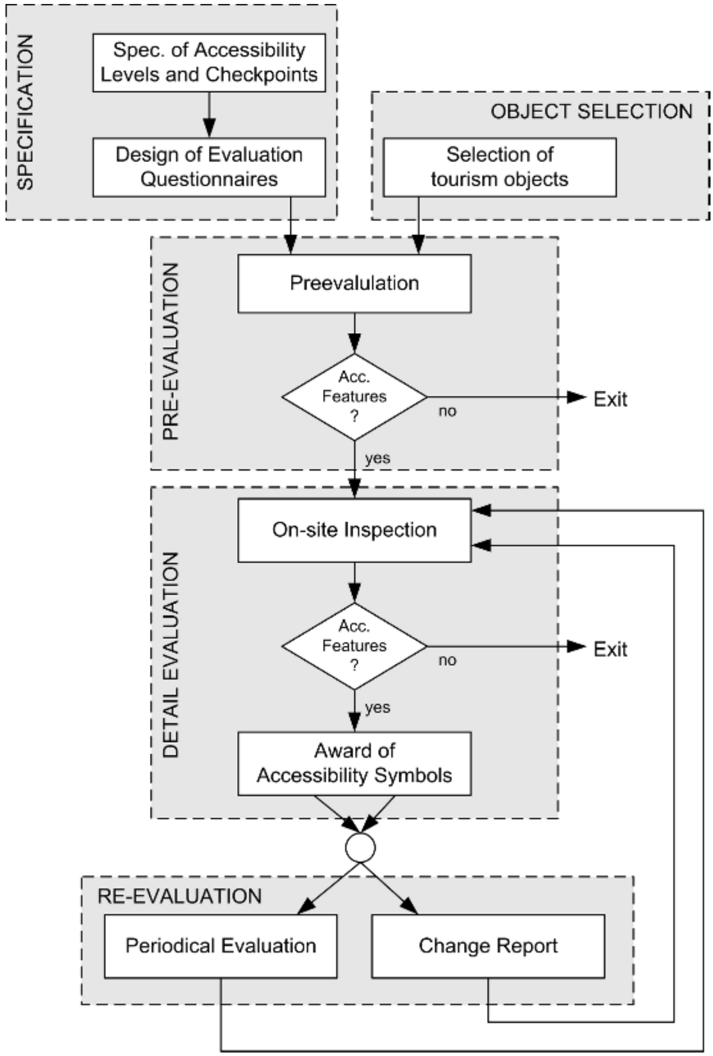


Fig. 2. Evaluation workflow

The steps "Specification" and "Object Selection" can be performed parallel, while the other three steps have to be processed as sequential steps.

Specification. During specification, the accessibility criterions for different object categories like hotels, restaurants, museums, etc. will be defined. All these criterions lead to a number of checkpoints, where each checkpoint has a certain priority and is assigned to a certain accessibility level. Together, all this checkpoints form the framework to evaluate accessibility and to map the evaluation results to the proposed accessibility label. To do pre- and detail evaluation, different questionnaires have to be implemented that reflect these checkpoints and enable the accessibility evaluation of tourism objects [10].

The specification of these criterions, the definition of its priority and the assignment to accessibility labels has to be done in close cooperation with people with disabilities and/or official representatives of all concerned disabilities.

Object Selection. Object selection summarizes the process to select tourism and infrastructure objects for evaluation. This can be applied in cooperation with tourism information provider, local authorities or on demand by providers of tourism objects.

Pre-Evaluation. Due to the fact, that on-site evaluation is a time consuming and costly activity. Pre-Evaluation has the goal to reduce the huge amount of tourism objects to a number of objects that offer at least a basic accessibility for one kind of disability. The pre-evaluation is implemented as a short questionnaire that will be sent to tourism object provider and can be filled out online or paper-based. Furthermore the pre-evaluation will reduce tourism object provider to the group that is interested to be evaluated and to use accessibility labels to attract the market share of people with disabilities.

Detail Evaluation. The detail evaluation is an on-site inspection that will be carried out by experts and specially trained evaluators. The evaluation will be carried out with questionnaires defined during the specification phase. After the evaluation the questionnaire will be analyzed, the checkpoints will be controlled and the affiliated accessibility labels will be awarded. The accessibility labels are valid up to the next detail evaluation.

Re-Evaluation. Up-to-dateness is one of the most important characteristics in information presentation. Accessibility characteristics have to be evaluated after each change. To fulfill this requirement, the tourism object will be re-evaluated in detail after the tourism object provider has reported a change. If there is no change report from the object provider, the tourism object will be re-evaluated in regular intervals (e.g.: 2 or 3 years).

Applying this evaluation process enables the evaluation of accessibility features for tourism objects and to award standardized accessibility labels that represent the accessibility for different kinds of accessibility on different levels. These accessibility labels can be used to promote the accessibility features of the tourism object on Webpages or information materials.

4 Conclusion

Accessible tourism creates competitive advantages, particularly from an economic point of view. People with disabilities together with elderly people share the same market segment. Because of the increasing share of older people the market segment is steadily increasing.

The needs of people with disabilities and elderly people are diverse. The design of services, products and offers, as well as the information presentation must reflect these requirements. To attract the market segment, the tourism industry has to understand the needs of these customers and must be consistently implemented in the context of an overall plan. To improve the quality of services for people with disabilities and to cope with their needs the tourism industry has to support tourists in an appropriate way in all three stages described. Those, who recognize accessible tourism as a quality criterion and as a competitive advantage first, will be the winners.

On the part of the tourists all users will profit from an improved infrastructure, extended offers and a more detailed and precise information offer. Frequently a prerequisite for disabled tourists is an increase of comfort and quality for other tourists, what contributes to a more pleasant stay and therefore to a qualitatively better vacation experience.

References

1. European Disability Forum (EDF), EDF Position Paper on Tourism: Framing the Future of European Tourism, Doc. EDF 01/13 EN, Brussels, Belgium (2001)
2. Neumann, P., Reuber, P., et al., Economic Impulses of Accessible Tourism for All, Study commissioned by the Federal Ministry of Economics and Labour (BMWA), ISSN 0342-9288, Berlin, Germany (2004)
3. National Disability Authority Udaras Naisiunta Michumais (NDA), Accessibility and Tourism – Submission to Tourism Policy Review Group, Dublin, Irland (2003)
4. Berdel, D., Gödl, D., Schoibl, H., Qualitätskriterien im Tourismus für Behinderte und ältere Menschen, Studie im Auftrag des Bundesministeriums für soziale Sicherheit und Generationen, Vienna, Austria (2002)
5. Hitsch, W., Probleme, Risken und Chancen des barrierefreien Tourismus, Diplomarbeit am Institut für Unternehmensführung, Tourismus und Dienstleistungswirtschaft der Leopold-Franzens-Universität, Innsbruck,Austria (2005)
6. Pühretmair, F., It's time to make eTourism accessible, 9th International Conference on Computers Helping People with Special Needs (ICCHP 2004), ISBN 3-540-22334-7, Springer Verlag, pp. 272-279, Paris, France (2004)
7. Oertel, B., Hasse, C., Scheermesser, M., Thio, S. L., Feil, T., Accessibility of Tourism Web sites within the European Union, Proceedings of the 11th International Conference on Information and Communication Technologies in Tourism (ENTER 2004), ISBN 3-211-20669-8, Springer Verlag, pp. 358 - 368, Cairo, Egypt (2004)
8. World Wide Web Consortium, Web Accessibility Initiative (WAI), Web Content Accessibility Guidelines (WCAG), online available on `http://www.w3c.org/wai` (25.01.2006)
9. Pühretmair, F., Miesenberger, K. - Making sense of accessibility in IT Design - usable accessibility vs. accessible usability, Sixteenth International Workshop on Database and Expert Systems Applications (DEXA 2005), ISBN 0-7695-2424-9, IEEE Computer Society, pp. 861-865, Copenhagen, Denmark (2005)
10. IHA – Hotelverband Deutschland, Marktsegment – Barrierefreie Hotels, Zielvereinbarungen zur Barrierefreiheit im Gastgewerbe, Checklisten zur Überprüfung der Mindeststandards in Beherbergungs- und Gastronomiebetrieben, online available on `http://www.hotellerie.de/home/barrierefreie_hotels_1859_1856.html` (23.03.2006)

People with Disabilities: Materials for Teaching Accessibility and Design for All

Introduction to the Special Thematic Session

Gerhard Weber[1] and Julio Abascal[2]

[1] Multimedia Campus Kiel, Human-Centered Interfaces Research Group
Westring 431-451, 24118 Kiel, Germany
g.weber@mmc-kiel.com

[2] Laboratory of HCI for Special Needs, University of the Basque Country
Donostia-San Sebastian, Spain
julio.abascal@ehu.es

Abstract. The growing interest on teaching accessibility and design for all requires good teaching material. While the research body is growing, emphasis for teaching materials is less developed. We identify a few criteria for teaching materials and identify their role in curriculum and course development.

1 Introduction

Teaching materials on accessibility and design for all are only slowly becoming available. While edited books [6,9,11] are giving an overview on different but considerable large aspects of the field, most interesting are annual or bi-annual conference proceedings such as from AAATE, ASSETS, CVHI, CUU, ICCHP, UAHCI to name a few. Together with journal publications the number of references is considerably large. It appears to be difficult to expect from lecturers to consult this body of research when teaching.

Likewise, standards such as ISO 9241 (-171) are not addressing beginners and appear to be too abstract. Their mapping into guidelines may be more suitable for purposes of teaching and learning. The most prominent example of such guidelines, the Web content accessibility guidelines [10], have shown that their application is very well linked the ability to learn them [2]. There is no shortage of general information on web accessibility but not enough teaching material available to help students of different disciplines such as business studies, design, electrical engineering or computer science to learn how to identify which design decisions on business process, devices or software have an impact on a particular group of users.

A common request by lecturers is aiming at case studies, similar to the role of teaching materials in general HCI education. For example, hearing a screen reader in action can explain how access by blind people to the web is working [4]. In addition, visualization of spoken interaction is required to deepen the understanding of the limitations a screen reader user experiences when browsing the web. Such a visualization may require the viewer to think being the first person (the user with disability) or being a neutral observer. In some areas, including haptic interaction,

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 337 – 340, 2006.

visualization is still to be developed and only direct experience can help students to understand, for example, guidelines on tactile graphics. Still, it is demanding for a management student to develop from spoken access to an eShops the ability to request an accessible form and prepare the proper criteria to compare different solutions.

In the following we will explore the role of training materials to mature the field of accessibility and design for all from the point of view of integrating it into higher education, and hence curriculum and course development. As several authors in this session have noted, plays the web an essential role to share training material.

2 Situating Design for All in Higher Education

Research work in the area of accessibility and *design for all* can develop unique results or is integrating previous work while extending it to comply with extraordinary user needs. Acknowledgement of this research work is an important aspect to understand the importance of the individual achievement. Recognitions such as the ACM Best Dissertation Award granted to T. V. Raman and the Social Impact Award of ACM SIGCHI in 2006 to Ted Henter help to promote the field. Recognition of web designer's success in making web pages more accessible by the stakeholders in the area of web accessibility may reverse the situation and attracts more practitioners to learn from researchers. The BIENE award is a good example for attracting a large audience of professionals [1].

However, students aiming at an educational degree in the area of special needs education, for example, are expected to have a very general overview about assistive devices and about the applicable pedagogic methods associated for example to sign language of the Braille notation. The learning objective accessibility implies most of all to study factual knowledge [3, 8] for a degree in design, engineering or computer science. In order to apply design for all principles a more procedural knowledge is required to ensure a methodological competence in analyzing the context of use [7, 11].

Assessment of student's success in mastering factual knowledge is following the same principles as for other knowledge areas. Quizzes with multiple choice questions are a quick way to pinpoint some issues. Lab work may be required to monitor progress in understanding the needs of users. More comprehensive observation is possible if a student writes a report, analyses existing work based on the reading list assigned to a given course and applies the concept to a novel paradigm such as AJAX-based communication between user interfaces in web applications and back-end systems.

Training materials support the learning objective if students can develop into young researchers. For example, student projects on accessible tables may well lead into novel browser paradigms of simplifying table navigation through transformation, better document standards such as for mobile phones or auditory interfaces to tabular data.

Quality of teaching material therefore can be identified by checking the following criteria:

1. Essential factual knowledge is covered to understand particular user group's needs.
2. Process or method is covered in order to identify and apply a user group's needs.
3. Knowledge covered by training material can be assessed.
4. Individual learning objective can lead to further courses and contributes to curriculum.

3 Curriculum Development

It is evident that Inclusive Design has grown in importance in the last years, becoming of vital significance in some specific areas, such as Informatics and Telematics studies. Experts consider that future designers should know and master the main concepts of the design for all. This means that these concepts should be included in all graduate studies with a design component. In addition, it is necessary to define the structure of graduate studies focused on Inclusive Design. Among the diverse initiatives that have been taken to fill this need IDCnet stands out.

IDCnet was a Thematic Network, supported by the Information Society Technologies Programme of the European Commission, devoted to the development of a curriculum in Inclusive Design. Although the project officially finished in 2005 its activities continue under the umbrella of the European Design for All e-Accessibility Network (EDeAN).

IDCnet held a workshop in Helsinki (Finland) in February of 2003 on "Design for All Curriculum: Towards a synergy of the needs of ICT industry and education" that gathered experts from industry and academia, as well as from the European Commission. The second workshop was held in Sankt Augustin (Germany) in January of 2004, focused on the topic "DfA Education & Research Policies and Strategies", intended to support the development of recommendations for Design for All education and research policies and strategies in Europe.

IDCnet produced many interesting deliverables that can be used as a starting point to integrate Inclusive Design in higher education [5]. To this end, the following documents are highly recommendable:

- D2.2 - The optimal graduate profile for DfA based on the needs of industry and the possibilities of within educational institutions
- D3.2 - Identifying Core Knowledge and Skill Sets for Model Curricula
- D3.3 - Teaching DfA Core Knowledge and Skill Sets
- D4.2 - Assessment of and recommendations on DfA related higher education and research policies and strategies in EU countries

4 Course Development

Usually, people aiming to include *design for all* in undergraduate courses have to develop their own teaching materials (texts, practical exercises, etc.) from the scratch. This considerable effort may discourage many lecturers that frequently have not enough support to develop this task. A good solution for this problem is the sharing of teaching materials. In this way each lecturer can profit from the successful experiences and tested materials developed by other people and can also contribute with their own materials. Shared materials can include texts, presentations, laboratories, examples of good practices and guidelines to use them.

Internet provides now-a-days excellent support for sharing of teaching materials allowing the creation of repositories of well structured and commented documents. Such a structure enhances the added value of the common resources, providing complementary information on when, how, where to use the diverse materials.

This special thematic session on Materials for Teaching Accessibility and Design for All aims to discus diverse approaches to the implementation of Inclusive Design courses and the possibility of sharing teaching materials to easy the work of pioneer lecturers.

This activity is supported by the Working Group 13.3 on HCI and Disability of the International federation for Information Processing. One of the principal objectives of the working group is to make HCI designers aware of the needs of people with disabilities. IFIP WG13.3 has organized successive workshops to facilitate the advancement in the use of Guidelines for Inclusive Design and published related materials [7].

References

1. http://www.biene-award.de, retrieved April 11, 2006
2. Disability Rights Commission. (2004). *The web: access and inclusion for disabled people.* London: The Stationery Office
3. Edwards, A; Edwards, A, and Mynatt, E. (2003) Designing for Users with Special Needs, CHI 2003 Tutorial 29
4. European Commission (2006) Web Sites That Work, http://www.idcnet.info/wai_video, retrieved 25. January 2006
5. IDCnet Deliverables: Inclusive Design Curriculum Network. http://www.idcnet.info/ documents
6. Lazar, J. (2006) Universal usability, Wiley
7. Nicolle C., Abascal J. (Eds.) (2001) Inclusive Design Guidelines for HCI. Taylor & Francis. London
8. Stephanidis, C.; Savidis, A. and Akoumianakis, D (2001) Developing Universally Accessible User Interfaces: Concepts, Methods and Tools, CHI 2001 tutorial 28
9. Stephanidis, C. (ed.) (2001) User interfaces for all - Concepts, Methods, and Tools, Lawrence Erlbaum
10. Web content accessibility guidelines 1.0 (1999) http://www.w3.org/TR/WCAG10/
11. Welch, P. (1995). Strategies for Teaching Universal Design, Adaptive Environments Center, 1995, also http://www.adaptiveenvironments.org/universal/strategies.phpï

Teaching Web Accessibility with "Contramano" and Hera

Carlos Benavídez[1], José L. Fuertes[2], Emmanuelle Gutiérrez[1], and Loïc Martínez[2]

[1] Sidar Foundation. Spain
{carlos, emmanuelle}@sidar.org
[2] Universidad Politécnica de Madrid. Spain
{jfuertes, loic}@fi.upm.es

Abstract. There is a need for training in design for all both at universities and in organisations, particularly as regards accessible web design. In this paper we present the experiences of the Sidar Foundation and the Technical University of Madrid in teaching web accessibility, focusing on the use of two key materials. The first one is Contramano, a fictitious web site designed to fail every accessibility checkpoint. It can be used to give examples of bad practices or as a basis for short exercises focused on specific checkpoints. The second material is the HERA tool, an on-line evaluation tool that has been designed to assist the manual evaluation of web accessibility. The paper will present the experience of using these materials in both higher education and other courses.

1 Introduction

In recent years the design for all concept has been increasingly promoted to deal with accessibility issues in both the physical world and, of course, the so-called Information Society [1,2].

In this context there is the common perception that education has a key role to play in the future of an inclusive society for all: the design for all concepts, principles and techniques should be introduced in the curricula of all occupations working on the built environment, including architects, civil engineers and software engineers [3,4]. In the case of Information and Communication Technologies (ICT) there are several organisations and projects working on the development of curricula and materials for design for all training, like IDCnet [5], EDeAN [6] and IFIP WG 13.3 [7].

In this paper we present the two most important materials that the Sidar Foundation [8] and the Technical University of Madrid (UPM) have been using to teach accessible web design in courses inside and outside the university: Contramano, a fictitious web site, and HERA, a web accessibility evaluation tool.

The paper is organised as follows. Section 2 will present the Contramano web site and its additional teaching material. Section 3 will briefly describe the HERA tool and its main educational features. Section 4 will discuss our experiences in the use of Contramano and HERA when teaching accessible web design. Finally, section 5 will present some concluding remarks, along with work to be undertaken in the future.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 341 – 348, 2006.

2 Contramano

Contramano [9] is a fictitious web site for left-handed people, offering information, games, curiosities and an electronic shop of specially created goods. The name Contramano comes from the Spanish expression “a contramano”, which means “going the wrong way” and is used in some South-American countries as a synonym of “moving on the left”.

This web site is useful for both teachers and students. It is a comprehensive web site for students to review and modify, using which they can do short practical exercises on accessible web evaluation and design. Also it is a complete collection of bad practices giving simple examples for all of the Web Content Accessibility Guidelines (WCAG) checkpoints [10], techniques to be avoided, and typical accessibility barriers.

As teaching material, Contramano consists of an incorrect and a correct web site, a collection of exercises and a teacher’s guide.

2.1 Contramano Incorrect Web Site

The incorrect version of Contramano consists of a web site full of accessibility and usability mistakes. This web site is divided into eight sections:

- Initial page - A “splash-screen” flash animation.
- Main page - A navigational page offering links to the sections of the web site.
- Presentation - It contains introductory content about being left-handed and mini-games related to palindromes.
- Myths and legends - It contains several myths and legends related to the left-handed and a link to a bilingual book description.
- The brain - It explains the organisation of and differences between the brains of left- and right-handed people. There is a server-side map explaining the main parts of the brain.
- Curiosities - It contains funny facts and statistics about the use of the left hand and foot. There is also an animation of the word “contramano” presented in finger spelling and a mini-game related to a famous old painting.
- Links - A set of links to other left-handed web sites and a practical joke.
- People - List of famous left-handed people, including musicians, politicians, writers and sportsmen. There is a video of a well-known goal scored by Maradona.
- Shopping - This is the commercial part of the web site, containing several products for left-handed people: t-shirts, writing instruments, watches and other utensils. In addition it contains information about the international day of left-handed people.

Each of the above sections and pages was designed with the goal of violating some of the WCAG checkpoints, allowing the teacher to show examples of bad practices and enabling the student to review and correct those problems. Table 1 contains an overview of the educational goals and checkpoint coverage of each of the Contramano sections. It should be noted that all the pages of the incorrect version of the web site fail to pass some checkpoints, like 3.2 (conform to formal grammars) and 4.3 (identify the primary natural language).

Table 1. Educational goals and checkpoint coverage of the Contramano web site

Section	Main goals	Priority 1	Priority 2	Priority 3
Initial page	Accessible inclusion of objects. Bad practice: an animation as start page.	1.1, 6.3, 8.1, 11.4	3.2, 6.4, 6.5, 9.3	4.3, 11.3
Main page	Alternative texts and long descriptions. Colour contrast. Non-relative font size. Use of quotes. Layout tables. Identification of link targets. Metadata. Consistent navigation and presentation style.	1.1, 14.1	2.2, 3.2, 3.3, 3.4, 3.7, 5.3, 5.4, 11.2, 13.1, 13.2, 13.4, 13.6	2.2, 4.3, 9.4, 13.5, 14.2, 14.3
Presentation	Colour contrast. Use of images instead of text. Header elements. Natural language changes. JavaScript-only link target. Physical events. Time-dependent interaction. Tab-index behaviour.	4.1, 6.3, 14.1	2.2, 3.1, 3.2, 3.5, 7.2, 6.4, 7.2, 7.4, 9.3, 13.1, 13.2, 13.5	2.2, 4.3, 9.4, 13.5, 14.2, 14.3
Myths and legends	Alternative text. Use of list. Header elements. Natural language changes. Pop-up windows. Tables with side-by-side text.	1.1, 4.1	3.2, 3.5, 3.6, 3.7, 10.1	4.3, 10.3
The brain	Alternative texts. Images for layout. Pop-up windows. Server-side image map. Non use of style sheets.	1.1, 1.2, 9.1	3.2, 3.3, 3.5, 13.1, 13.4, 13.5	4.3, 13.5, 14.2, 14.3
Curiosities	Non accessible flash object. Use of lists. Abbreviations and acronyms. Moving text. Loss of structure when style sheets are turned off. Obsolete elements.	1.1, 6.1	3.2, 3.6, 3.7, 7.3, 7.5, 8.1, 9.2, 11.1, 11.2, 12.3, 13.2, 13.4	4.2, 4.3, 13.5, 14.2, 14.3
Links	Relevant link text. ASCII-art. Relevant information at the beginning of list items.	1.1	3.2, 13.1	4.3, 13.8, 13.10
People	Non accessible video. Two-column layout table. Use of lists.	1.1, 1.3, 1.4	3.2, 3.3, 3.6	4.3
Shopping	Use of frames. Use of data tables. Colour-only information. Client image maps. Non-separated consecutive links. Search functionality. Screen flickering. Moving images.	1.1, 2.1, 5.2, 6.2, 6.3, 7.1, 9.1, 12.1	3.1, 3.2, 5.3, 5.4, 6.4, 6.5, 9.2, 9.3, 10.1, 10.2, 12.2, 12.3, 12.4, 13.2	1.5, 4.3, 5.6, 9.5, 10.4, 10.5, 13.6, 13.7, 13.9, 14.2

2.2 Contramano Correct Web Site

We have also developed a correct, accessible version of the web site, keeping as much of the original aesthetics as possible.

The accessibility problems of the original site have been solved in the correct version. For instance, the initial page now contains more content than just the flash animation, the contents are better structured, the navigation and presentation style is consistent across the pages, etc.

The main goal of the correct version is to enable the teacher to quickly show how to remove the accessibility barriers there are in the original site. Consequently, the correct version of Contramano does not aim to be a visually attractive and fashionable web site, but a site conforming to the WCAG.

2.3 Exercises

Accompanying the Contramano incorrect and correct versions, there is a set of 67 exercises corresponding to examples of bad practices in the application of the WCAG checkpoints.

These exercises can be navigated in three ways: by the sorting order of the related checkpoints, by the type of elements they refer to or by priority order. This way students can follow the order proposed by the teacher or their own preferred order when working on their own outside the classroom.

The exercises consist of small fragments of the Contramano web site, each of one focused on only one of the WCAG checkpoints and accompanied by the required files (HTML, images, videos, etc.). In this way the student only has to worry about the checkpoint the exercise deals with.

The exercises are presented as web pages, with two different presentation styles, the aim being here to accommodate a wider range of users. These pages contain the element up for modification, along with links to the needed files and to the original content in the Contramano site.

2.4 Teacher's Guide

Teachers of web accessibility courses are experts with extensive knowledge of WCAG and its application. However, even an expert will not always remember the full text of each of the checkpoints and, furthermore, will not always recall which pages of Contramano provide examples for each of the checkpoints.

For this reason, we have developed a teacher's guide accompanying the Contramano site. This guide contains, for each of the checkpoints, its identifier, its priority level, its full text and the failure examples in Contramano (with links to the respective page). The checkpoints can be navigated in the same orders as described for the exercises.

In addition, the guide contains specific help about evaluation techniques for each of the checkpoints. This information has been gathered by the Sidar Foundation through years of experience in the evaluation of web site accessibility.

Finally, the teacher's guide provides a mini-application for the management of the information it contains. Thus, teachers can modify the guide's content, adding more relationships between the checkpoints and the respective examples or exercises.

3 HERA

HERA is an on-line tool that provides support for the manual evaluation of the accessibility of web pages based on WCAG. HERA 1.0 was made freely available to the public in 2003 and was a great success, as has been demonstrated by the positive feedback received from system users. It was the first on-line tool for accessibility assessment that was developed with the manual evaluation process in mind [11]. HERA 2.0 was launched in 2005 to overcome some weaknesses of the previous version and to improve system usability [12].

HERA provides a full set of useful functionalities from the point of view of web accessibility education, which are summarised below (Figure 1):

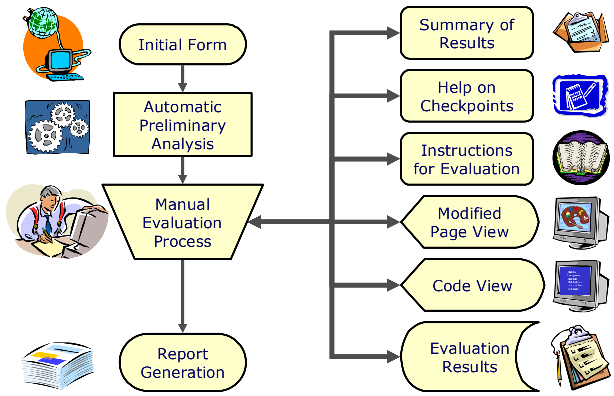


Fig. 1. Overview of HERA

Automatic Preliminary Analysis. HERA inspects the web page and automatically assigns a value for each of the 65 checkpoints: pass, fail, not applicable or needs checking. Table 2 shows a summary of how checkpoints are automatically evaluated with HERA. In this table several checkpoints appear in more than one column. For instance, checkpoint 1.1 (text alternatives) can be automatically evaluated as fail if there are images without the "alt" attribute; as non applicable if there are no non-textual elements (images, objects...) in the page; and as needs checking if there are images with "alt" attributes that require human evaluation to assess if the alternative text matches the image. On the other hand, this checkpoint will never be automatically evaluated as pass, because the computer cannot judge the adequacy of alternative texts.

Table 2. Summary of checkpoints automatically analysed by HERA

	Pass	Fail	N/A	Human
Priority 1	2	4	12	16
Priority 2	17	18	9	21
Priority 3	6	6	6	13
Total	**25**	**28**	**27**	**50**

Manual Evaluation Process. After the preliminary analysis, the evaluator has to proceed with the manual evaluation, using different navigation strategies. For instance, some evaluators follow the priority levels while others prefer to first focus on the failed checkpoints (irrespective of their priority level), then on checkpoints requiring manual evaluation and then finish with passed and non-applicable checkpoints. During the manual evaluation, users receive support from the other HERA modules.

Summary of Results. Once the preliminary analysis has been completed, HERA shows the results of this process, with information such as the URL under analysis, time spent on automatic evaluation, number of errors found and, more importantly, a table with the summary of the checkpoint evaluation results. This table shows, for each priority level, the number of checkpoints that require manual evaluation, that have passed, that have failed and that are not applicable. This table provides another navigational tool for the manual evaluation process and updates its content as the manual revision progresses.

Help on Checkpoints. This tool provides help about the meaning of each of the checkpoints, giving the full text of the checkpoint that is being checked. This is extremely useful in educational contexts, where the users have little knowledge of the WCAG.

Instructions for Evaluation. If required, HERA provides additional information about the checkpoint goal, techniques for its evaluation and techniques for a correct implementation of the checkpoint. Again this feature is relevant in an educational context, helping the novice evaluators to gain insight about the process of web accessibility evaluation.

Modified Page View. HERA can generate a modified view of a page, used to highlight with boxes, colours and icons the elements of the pages that have to be analysed for each checkpoint. This greatly minimises the need to examine the source code of the web page.

Code View. HERA can also show the source code of the page highlighting the items which should be checked, using the same colour scheme and icons that the modified page view, thus facilitating the identification of the parts of the source code that require additional inspection or modification.

Evaluation Results. HERA provides a service for the users to store the complete or incomplete evaluation results for further reference.

Report Generation. HERA provides a report generation module. The users can provide contextual information for the report (name and email of the evaluator, title of the report and a short commentary). Then the users can choose which of the checkpoints will appear in the generated report. The report can be generated in three different formats: XHTML, PDF or RDF (using the EARL vocabulary [13]).

4 Our Experience with Contramano and HERA

We have used both Contramano and HERA as part of several web accessibility courses inside and outside the university with different goals but the same satisfying results.

The courses taught at the UPM in collaboration with the Sidar Foundation typically involve around 50 students and last one semester. These are courses with a low number of lectures spread across the semester, and students have to complete a practical exercise on the design and implementation of a small-scale accessible web site.

In these courses we use Contramano during the lectures to illustrate examples of bad practices for the more relevant WCAG checkpoints. And we use HERA as the compulsory tool for students to evaluate their small web site and to generate the respective accessibility report that they have to present as part of their practical work.

The courses taught by the Sidar Foundation outside the university (with the collaboration of teachers from the UPM) are completely different. Typically they are 20-hour 3- or 4-day intensive courses for up to 20 students in computer classrooms.

In these short courses we use Contramano and HERA differently. After some introductory material, we start using HERA early on, the aim being for students to get a better knowledge of the typical accessibility issues of web sites. Then we proceed with the WCAG checkpoints, grouped around the main elements. During this explanation we use Contramano to illustrate examples of bad practices and to set short exercises to be completed by the students in the classroom.

In either case we have found the use of Contramano and HERA to be extremely beneficial for both teachers and students. We have always been awarded good scores in the course evaluation surveys that the students fill in at the end of the courses.

5 Conclusions and Future Work

A lot of emphasis is being placed on accessibility and design for all in all domains today, including the Information Society. Education in design for all, where materials and examples of good practices are being collected, is a key issue if an inclusive society is to materialise.

In this paper we have presented two teaching materials developed by the Sidar Foundation, in collaboration with the UPM: Contramano, a fictitious web site, and HERA, a web accessibility evaluation tool.

These materials have been successfully used by the Sidar Foundation and UPM in several courses on web accessibility taught inside and outside the University.

However, there is always room for improvement. Here are some ideas for future work:

- The design and contents of the non-accessible version of Contramano could be updated to better reflect the typical accessibility mistakes occurring today.
- The accessible version of Contramano could be redesigned, mainly to improve its aesthetics and make it more appealing as a web site.
- Contramano could also be translated to other languages (now it is only written in Spanish), so it could be use for international courses.
- Finally, HERA 2.0 should enable the user to navigate the checkpoints in the same ordering schemata as they are presented in the Contramano exercises, easing the combined use of the two educational materials. This is something planned for the final version of HERA 2.0.

References

1. Political Declaration of the II World Summit of Cities and Local Authorities on the Information Society. Bilbao, Spain. November, 11 (2005) (URL: `http://www.it4all-bilbao.org/declaracion`).
2. Tunis Commitment. World Summit on the Information Society. Document WSIS-05/TUNIS/DOC/7-E. Tunis. November, 18. (2005) (URL: `http://www.itu.int/wsis/docs2/tunis/off/7.html`).

3. Council of Europe. Resolution ResAP(2001)1. On the introduction of the principles of universal design into the curricula of all occupations working on the built environment (Adopted by the Committee of Ministers on 15 February 2001, at the 742nd meeting of the Ministers Deputies). (2001).
4. Spanish Official State Bulletin (BOE). Law 51/2003 of Equal Opportunities, Non-Discrimination and Universal Access for Persons with Disabilities. Tenth final Disposition: Curricula on Universal Accessibility and Professional Formation (2003).
5. Inclusive Design Curriculum Network (IDCnet). Homepage. (2005) (URL: http://www.idcnet.info/desc).
6. European Design for All e-Accessibility Network (EDeAN). Homepage. (2006) (URL: http://www.eaccessibility.org/).
7. International Federation for Information Processing (IFIP). Working Group 13.3 Human Computer Interaction and Disabilities. Homepage. (2005) (URL: http://www.info.fundp.ac.be/IFIP13-3/).
8. Sidar Foundation. Homepage (in Spanish) (2006) (URL: http://www.sidar.org/).
9. Sidar Foundation. Contramano presentation (in Spanish) (2005). (URL: http://www.sidar.org/contramano/).
10. Chisholm, W., Vanderheiden, G., Jacobs, I.: Web Content Accessibility Guidelines 1.0. World Wide Web Consortium Recommendation (1999) (URL: http://www.w3.org/TR/ WCAG10).
11. Benavidez, C., Gutiérrez, E.: HERA una herramienta para la revisión manual de la accesibilidad web. In: Jornadas de Accesibilidad y Nuevas Tecnologías (JANT 2004). Bilbao, Spain (2004) (in Spanish).
12. Benavidez, C. Fuertes, J. L., Gutiérrez, E., Martínez L., HERA: an on-line tool for the semi-automatic evaluation of web accessibility. In: Proceedings of the Euro American Conference on Telematics and Information Systems, EATIS 2006. Santa Marta, Colombia (2006).
13. McCathieNevile, C., Abou-Zahra, S.: Evaluation and Report Language (EARL) 1.0 Schema. World Wide Web Consortium (2005) (URL: http://www.w3.org/TR/EARL10/).

Introducing Media Managers to Usability and Accessibility

Christoph Haffner and Gerhard Weber

Multimedia Campus Kiel, Human-Centered Interfaces Research Group
Westring 431-451, 24118 Kiel, Germany
{c.haffner, g.weber}@mmc-kiel.com

Abstract. The degree program for M.Sc. Multimedia Management includes teaching on basic aspects of usability and accessibility. Our approach is to introduce accessibility issues as an application of a more general learning objective. Learners experience accessibility issues indirectly when working on a project involving multiple students. Two case studies are presented: a) developing an accessible Flash-based web interface and b) the heuristic evaluation of the usability of a novel software package. 17 students participated in a questionnaire allowing to assess the effectiveness of our approach.

1 Introduction

Accessibility is a major phenomenon which receives some attention as it is expected that better accessibility leads to more usable interactive systems. However, teaching about universal design [1] and users with special needs is reaching its limits if students have not experienced the limitations caused by the lack of communication with and through computers. An initial attempt to overcome such lack of insight through lack of own experience is to show a video. Several different kind of people can be viewed in a short time and use of assistive devices can be observed by students [2]. However, we believe such factual information needs to be connected to the methods and facts suitable for the general field of HCI and taught as part of the overall curriculum [3]. If accessibility is seen as a separate field by students it remains to be a niche and will not be addressed in projects aiming at mainstream users.

The curriculum for the degree program Master of Science degree in Multimedia Management provides a unique opportunity to introduce accessibility to an international studentship, who for the majority has never heard about handicapped people using computers. Designing for disabled users is nearly impossible if the implementation of accessible interaction techniques is unknown. But implementation cannot be attempted without properly identifying user needs and designing some accessible systems. Such circular dependency can only be broken up if students are motivated for the instructional unit and feel responsible themselves to apply their knowledge.

Our approach to make students familiar with design for all principles is built on the following steps

- select a HCI curriculum topic which at first hand is not affected by accessibility issues,
- introduce design for all principles as an application of the chosen topic,

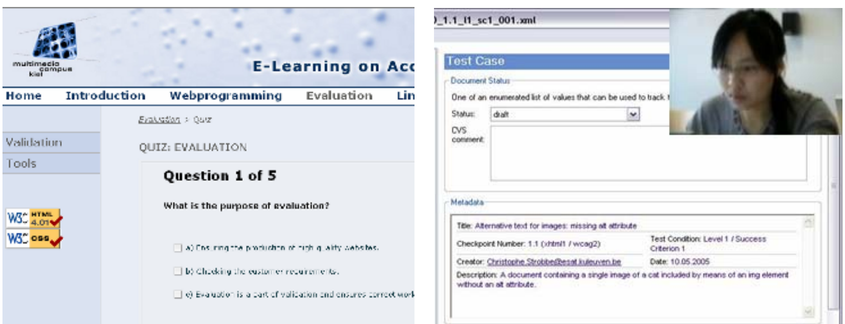
K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 349 – 352, 2006.

- allow students to learn the topic and apply their knowledge by relying on design for all methods, and
- perform an assessment.

Thereby universal design serves a double purpose. It is both a learning objective in its own and a HCI technique applicable to a particular HCI method. The introduction to evaluation methods used in HCI and the development of a web-based interactive system for assessments have served us as very different curriculum topics in the past two academic years.

2 Case Studies

In the following, two case studies shall illustrate the approach. Teaching of management principles is part of a team project, here students face the need to organize themselves and their team members in order to develop a software application (see Fig. 1 a). An HCI topic are empirical usability studies which include learning to apply multiple methods and tools (see Fig. 1 b).



a) b)

Fig. 1. a) Accessible quiz; b) recording a usability test scenario

2.1 E-Learning Unit on Accessibility

A case study was carried out within a student's team project [4]. Coming up with the title "E-Learning on Accessibility", the focus of the project was to implement an e-learning title to teach the main matters of accessibility in general and web accessibility in detail. The acquired knowledge was to be tested in a quiz. A market analysis studied the need for such eLearning titles. The implemented e-learning title was designed as a web application with a quiz implemented in accessible Flash. Following our approach the title had itself to be accessible according to WAI guidelines by considering "Level Double-A Conformance"[5].

This approach for teaching principles and guidelines of accessibility forced the team to collect information on accessibility issues and to develop the didactics for web-based eLearning concept. By applying that knowledge in a concrete implementation, the gained knowledge should be deepened sustainable.

2.2 Empirical Usability Studies

In the context of a computer graphics lecture, accessibility and usability are taught to sensitize future media managers for the usage of basic technologies and methods of that area. The theoretical topic, provided in the lecture was imparted as practical knowledge and experience in an exercise. The students evaluated a novel software tool called "TCE – Test Case Editor" by accomplishing user tests. The TCE tool is a new editor for systematic organization of accessibility evaluations [6]. TCE requires the user to become familiar with different types of disabilities and assistive technologies. The method applied for user testing was divided into the following parts:

- A "software expert" has to acquaint himself with the TCE tool in detail. He defines four tasks of different complexity a test user has to solve,
- two subjects without having previous knowledge of the TCE tool shall try to accomplish the tasks defined by the "Software Expert",
- a "video expert" records the computer screen, the subject's face and voice during the tasks and edits the audio and video content live to provide all information within one single screen,
- an "observer" watches the recordings and describes verbally the subjects acting and comments as well as noticeable attitudes. All observations including proceeding time are logged using the "Ovo Logger" Software tool [7]. Based on the observations, a final report is generated by the logging software. The "Observer" enters the users and tasks in a data base and allocates them to heuristic evaluation categories.

3 Findings

Case study one shows students can make themselves familiar with web accessibility in the given educational framework. While the students had to look into the subject in depth in theory and practice, the understanding of the needs of special user groups was correspondingly deep.

After finishing case study 2, a questionnaire was filled out by the students. The results of this evaluation help to measure the grade of the student's sensitization for the methods of usability and accessibility testing. Out of 17 respondents 65% were male, 35% were female. Also, 35% of students had learned before about usability issues. From those, 2 students had already professional experience in usability testing.

The evaluation's aim was to assess if the students recognized usability issues and distinguish them from accessibility issues. More then 70% agree on performing a usability test of the Test Case Editor for accessibility testing is a good approach to learn usability test methods and accessibility concepts within one single task. All students agreed that usability testing is an important step within the quality management of the software engineering process, actually 42% strongly agreed in that point. 76% agree about the statement, that accessibility provides the opportunity for business cases. In contrast, Figure 2 shows the response to the question if TCE is applicable to people with special needs. Students seems to lack a precise understanding of requirements of users with special needs vs. general usability issues. More training on assistive technologies is necessary.

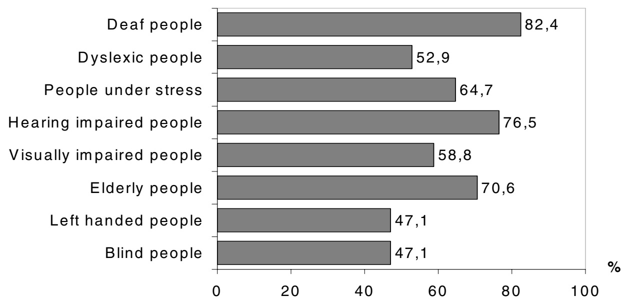


Fig. 2. Students identification of user groups

In the past, individual students who have become familiar with accessibility techniques also considered to intensify their knowledge and have developed their master thesis in this area. This rate is increased to 16% for the current academic year 2005/2006.

References

1. Welch, P. (1995). Strategies for Teaching Universal Design, Adaptive Environments Center, 1995, also
2. `http://www.adaptiveenvironments.org/universal/strategies.php`.
3. European Commission (2006) Web Sites That Work, `http://www.idcnet.info/wai_video`, retrieved 25. January 2006.
4. `http://www.idcnet.info/html/IDCnet_D3.2.1.html.`
5. Werner, A.-K.; Blome, P.; Wu, J, Hülskamp, D. (2005). E-Learning on Accessibility, `http://hci.mmc-kiel.com/accessibility/index.htm`.
6. Web content accessibility guidelines 1.0 (1999) http://www.w3.org/TR/WCAG10/
7. Herramhof, S.; Petrie, H.; Strobbe, C.; Vlachogiannis, E.; Weimann, K.; Weber, G. and Velasco, C.A. (2006). Test Case Management Tools for Accessibility Testing, in this volume.
8. `www.ovostudios.com`, retrieved 25. January 2006.

Evaluating the Length of Virtual Horizontal Bar Chart Columns Augmented with Wrench and Sound Feedback

Tatiana G. Evreinova, Grigori Evreinov, and Roope Raisamo

TAUCHI Computer-Human Interaction Unit
Department of Computer Sciences
FIN-33014 University of Tampere, Finland
{e_tg, grse, rr}@cs.uta.fi

Abstract. Augmented visualization of the mathematic and scientific data is an essential aid in training blind students' pre-calculus skills. Compared to existing multidimensional wrench-reflection interfaces, one-dimensional stylus-based interaction concept could support blind users with reasonable feedback in different tasks. We designed a mock-up of the cable-suspended haptic interface and a match game-like piece of software to investigate the perception features of the length of the virtual horizontal bar chart columns augmented with wrench and sound feedback. The performance of the eight blindfolded subjects was evaluated in terms of the number of repeated inspections to detect twin chart columns with similar length, and the task completion time required to perform the chart inspection. The experience acquired within simulated gameplay conditions with the use of implemented cable-suspended interface can be applied in developing novel didactic tools for training blind students in estimating linear dimensions of the simulated objects.

1 Introduction

Augmented visualization of the mathematic and scientific data, especially graphs and diagrams, is an essential aid in training blind students' pre-calculus skills. Several techniques are being used to assist blind students in accessing graphical information. Graphs and diagrams can be presented as the tactile embossed patterns accompanied with Braille and speech cues. However, only 26% of blind university students are expert Braille readers [8]. In addition, low perceptive accuracy significantly impedes the proper imagination of the embossed graphic information.

Sribunruangrit et al. [7] studied a possibility to display the virtual graphic images employing the "Braille Box" device to control an array of Braille pins (4 by 4) with the touch tablet. The blind user specified the region of interest of the virtual graphic image being partially displayed under the fingertip of the one hand by moving the stylus along the graphic tablet by the other hand.

However, the perceptual performance of the subjects in the tasks of tracking graphic images was low. The subjects experienced essential difficulties in cognitive integration of the sequential locations of the image features (details) as information about the location of the details within the image layout was insufficient. The

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 353 – 360, 2006.

general layout and details are being delivered through exploratory movements and tactual-kinesthetic relationships which both have to be strictly coordinated in space and time to provide a cross-linkage between details, their location and the entire image layout [4]. The number of parameters used to provide the information about the image features (details and location) should be reasonable, but not redundant [2].

There have been several attempts recently to explore stylus-based wrench-reflection haptic devices whose actuators provide a force and direction moment vector to a user's hand for haptic interaction with virtual objects especially line, graphs and bar charts. As a rule, haptic features are assigned to each vertex of a virtual object. Wrench-reflection haptic devices allow perceiving only a single vertex at a time and then the subject has to mentally integrate a series of the vertices explored [8]. Grabowski and Barner [1] investigated a possibility to use a six-degrees of freedom force feedback PHANToM haptic device having a 3D force feedback workspace of about 160 mm × 120 mm × 120 mm [6] to augment haptic visualization for blind and visually impaired persons. They showed that the users were only nearly able to evaluate the layout of the graph presented with the device. The visually impaired users could not properly use the gridlines to find the maximum and minimum points of the virtual lines as well as they were not able to distinguish correctly other haptic cues.

As a candidate to produce a flexible and non-expensive feedback a thin cable may be attached to a stylus. The length and a tension of the cable might be controllable by means of a DC-motor and a pulley which can produce a gradual reflection force moment in dependence on a stylus position doing a restriction in some predefined direction and location or/and to simulate surface features. Due to a special kinematics and the points of applied mechanical moments the cable-suspended interface has a good force-moment/power ratio. There were already several tension based haptic interfaces [3,5] developed to deliver force moment to the user's fingers thus allowing presenting the features of the virtual objects in either 2D or 3D space using a force feedback. Tension based haptic interfaces have a scalable workspace. Nevertheless, they are still cumbersome to use due to the great number of the cables and the way of their fixation on the fingers.

The goal of this work was to investigate the perception features of force feedback provided with simplified one-dimensional cable suspended haptic interface in the task of the blind inspection and matching of the length of the virtual horizontal bar chart columns augmented with wrench and sound feedback.

2 Method Design

2.1 Manipulandum

The mock-up of the cable-suspended interface comprised of AceCad AceCat Flair USB Graphics tablet and the haptic unit providing cable tension at the particular moment and location. The tablet has active manipulandum surface of 127 mm × 95.25 mm and includes cordless pen. The prototype of the haptic unit consists of a FF-130SH DC motor having shaft radius of 1 mm, pulley, and reduction gear, to transform with a minimum lost, the torque of about 6 mN/m into a tractive force lead to a stylus. The reduction ratio used was 21:1. DC motor working in a pulse wide modulation mode provided the resulting tractive force of about 6 N.

2.2 Apparatus

The match game testing software was developed in Microsoft Visual Basic 6.0 under Windows 2000. Each trial consisted of two phases: the preliminary inspection phase and confirmation phase. Doing the preliminary inspection, the subject followed the tractive force provided with cable-dragged stylus and explored the length of each column comprising the horizontal bar chart. The tractive power (P_{SD}) was proportional to the length of the chart columns and counted as follows:

$$P_{SD} = F_{SD} \times L \times \frac{T_1}{T_1 + T_2}, \tag{1}$$

where F_{SD}– the tractive force being applied to a stylus; L – the length of the chart column; T_1– the period of time when the DC motor was switched on; T_2– the period of time when the DC motor was switched off.

In the present study, the number of the columns in the horizontal bar chart was restricted to 4 though this number can be increased. When inspecting a column the subject followed the stylus along the column starting from the left border. The inspection was considered as completed when s/he had crossed the right border of the column. “Negative” sound indicated crossing the edges in the intermediate point of the column, which was considered as an error. Each column of the graph was additionally augmented with a sound marker (earcons: do, re mi, fa) in order to specify the location of each column within the chart.

In the confirmation phase, the subject could continue an inspection of the chart columns if needed however the number of repeated inspections to explore each column was limited. Upon completion of the chart inspection, the subject had to choose two columns with a similar length. Pointing was accomplished with the stylus and accompanied with earcons; each pointed column was selected by pressing the space bar.

2.3 Participants

Eight sighted right-handed paid volunteers (six females and two males) from the staff and students at the University of Tampere participated in the study. The ages of the subjects ranged from 25 to 50 years with the mean age of 34. The average computer experience of all participants was 7.5 years. All of them used computers on a daily basis, reporting 7 to 10 hours of usage per day. All subjects had tried out a force feedback joystick, a tactile mouse and a tactile gamepad once or twice. No participants reported any loss or intolerance of haptic sensation. Since the technique is intended for blind users, the test subjects were blindfolded and wore a mask to simulate blindness condition.

2.4 Procedure

The subjects were tested individually. Prior to data collection, they were instructed concerning the features of the developed software and the overall testing procedure. The subjects were advised to anchor the hand on the desk where the cable-suspended haptic interface was located. They were asked to leave the hand relaxed while grasping the stylus and to follow the force moments provided with the device without applying strong resistance to the wrench move. Participants were then allowed to play a "warm-up" game. After that, the participants could ask questions if they felt that some steps in the testing procedure required more clarification. The experimental setup is shown in Figure 1.

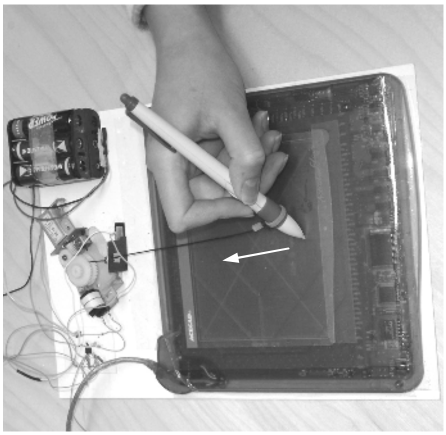


Fig. 1. The experimental setup: the subject anchors the hand on the desk when grasping a stylus and following the tractive force (white arrow on the figure) lead to a stylus to evaluate the length of the virtual horizontal bar chart columns

Each subject has completed three sessions during testing with no more than one session per day. During each test session, evaluating of chart with columns having both different minimum length and twin-column to non-twin column length ratio, was performed in a random order. The length of the chart column (L) was counted as follows:

$$L = L_{\min} \times k_{rnd} , \quad (2)$$

where $L_{\min}$ – minimum length of the column, k_{rnd} – a random value from 1 to 7.

The exploration of the chart was accomplished with chart columns having $L_{\min}$ value of 5, 10 and 13 mm. Seven different twin-column to non-twin column length ratios were varying on their values in dependency of $L_{\min}$ and k_{rnd} (Table 1).

Table 1. The parameters of the virtual horizontal bar chart columns used

Minimum length of the column, L_{min}					
5 mm		10 mm		13 mm	
Ratio, mm	Ratio reduced by cancellation	Ratio, mm	Ratio reduced by cancellation	Ratio, mm	Ratio reduced by cancellation
5:10	1:2	10:20	1:2	13:26	1:2
10:15	2:3	20:30	2:3	26:39	2:3
15:20	3:4	30:40	3:4	39:52	3:4
20:25	4:5	40:50	4:5	52:65	4:5
25:30	5:6	50:60	5:6	65:78	5:6
30:25	6:5	60:50	6:5	78:65	6:5
35:30	7:6	70:60	7:6	91:78	7:6

A session lasted for an hour in average. One test session consisted of three blocks. One test block consisted of 20 trials, that is, each subject performed 60 trials per one day, 180 trials in a total. This meant that subject had to successfully perform the confirmation phase for 20 times, i.e., to detect correctly the twin columns having similar length in each trial case. If it did not happen and the number of repeated inspections to explore four chart columns at confirmation phase exceeded the maximum number of repeated inspections allowed (30), the subject was automatically redirected to the inspection phase of the next chart. These data of incomplete trial were not stored in a log file and consequently excluded from further analysis.

The subjects could rest as desired between trials. Therefore, none of them reported loss or decrease of haptic sensation throughout the testing. The data concerning the behavioral patterns of the subjects' scanpaths at exploration of the virtual horizontal bar chart columns in the inspection phase were recorded and analyzed.

3 Results

The performance of the subjects maybe evaluated in terms of the number of repeated inspections to detect twin columns with similar length. The task completion time is comprised of the time spent to complete the preliminary inspection phase, the time spent to complete the confirmation phase that can also include repeated inspections, and the decision-making which of the chart columns were twins. As the number of columns was restricted to 4, most of the subjects had accomplished the thorough inspection during the first phase. We supposed that the time spent in this phase was the main part of the subjective performance; it had an impact on the final decision making process. Thus, the task completion time of the preliminary inspection phase is considered next as a parameter of the subject performance and was taken into further analysis. The behavioral patterns of the subjects' scanpaths made in the preliminary inspection phase were also analyzed. Some considerations regarding the average data recorded at evaluating the length of the virtual chart columns are summarized below.

3.1 Matching Twin Chart Columns

The average number of repeated inspections, made by the subjects to explore four horizontal bar chart columns in the preliminary inspection phase of the match game, is presented in Figure 2. The average number of repeated inspections varied from 1.5 to 3 when twin-column to non-twin column length ratio was less than 4:5 (see Table 1) and L_{min} was equal to 13 mm. The average number of repeated inspections increased when twin-column to non-twin column length ratio was more than 4:5 and varied from 2.31 to 3.75. The mean error rate recorded was of about 1.2%.

The number of repeated inspections varied from 4 to 6 when twin-column to non-twin column length ratio was less than 4:5 and L_{min} was equal to 10 mm. When twin-column to non-twin column length ratio was more than 4:5, the variation decreased to 2.25 – 4. The error rate increased up to 4.5% regarding the errors committed when L_{min} was 3 mm greater. The difference in the length of twin and non-twin chart columns became closer to the discrimination threshold. The performance of the subjects deteriorated respectively.

The number of repeated inspections varied from 3.7 to 6.5 when twin-column to non-twin column length ratio was less than 4:5 and L_{min} was equal to 5 mm, and increased to 4 - 7.5 when twin-column to non-twin column length ratio was more than 4:5. In this case, the difference in a length of twin and non-twin chart columns was small enough too or could probably be less than the discrimination threshold. Therefore, the performance of the subjects was extremely low. The mean error rate recorded was high, about 7.5%.

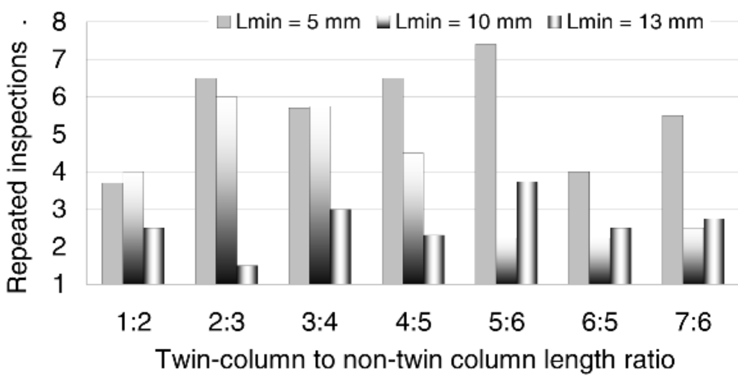


Fig. 2. The average number of repeated inspections made by the subjects to explore virtual horizontal bar chart columns in the preliminary inspection phase of the mach game

3.2 Task Completion Time

The mean twin-column detection time in the preliminary inspection phase of the match game is presented in Figure 3.

The mean task completion time varied from 4 to 8 s when twin-column to non-twin column length ratio was less than 4:5 (Table 1) and L_{min} was equal to 13 mm. When twin-column to non-twin column length ratio was more than 4:5, the mean task completion time slightly decreased and varied from 3 to 7 s.

In the inspection phase of the game the subjects needed from 7 to 12 s to complete the chart exploration when twin-column to non-twin column length ratio was less than 4:5 and L_{min} was equal to 10 mm. When twin-column to non-twin column length ratio was more than 4:5, the subjects have spent less time for the chart exploration but the range of time variation was increased to 6 – 14 s.

The subjects were required twice more time, of about 12 – 17 s, to explore the chart when twin-column to non-twin column length ratio was less than 4:5 and L_{min} was equal to 5 mm. The inspection of the chart took a greater time of about 7 – 17 s when twin-column to non-twin column length ratio was more than 4:5.

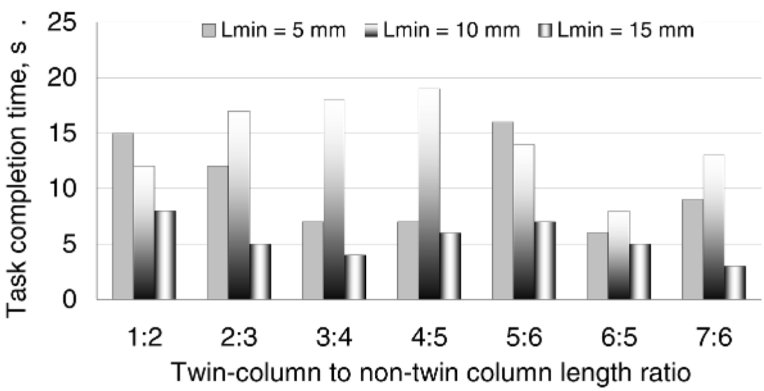


Fig. 3. The mean twin-column detection time at the preliminary inspection phase of the match game

3.3 Scanpaths Recorded

The subjects had different attitudes towards testing. Two of subjects were very careful with not making mistakes whereas others just wanted to complete the virtual chart inspection in the preliminary inspection phase of the match game as fast as possible without considering the increased number of errors at the confirmation phase. In the beginning of the testing, 6 of 8 subjects found sound feedback quite annoying although to the end of the experiment all the subjects were certain of that sound cues significantly facilitates the inspection procedure. The behavioral patterns confirmed that to evaluate the length of the virtual chart columns at the preliminary inspection phase 5 of 8 subjects had a tendency do not continue exploration of the rest part of the chart once upon a short chart exploration they were sure which columns were twins. 4 of 8 subjects explored all the chart columns one by one before they became sure which chart columns were twins. All the subjects were tended to explore all the chart columns one by one for a prolonged period of time in a case when the twin columns were weakly differed from each other.

4 Conclusion

We presented an evaluation of the simplified one-dimensional cable-suspended haptic interface and match game-like software in the task of blind inspection and matching of the length of the virtual horizontal bar chart columns augmented with wrench and sound feedback. The performance of the blindfolded subjects was evaluated in terms of the number of repeated inspections to detect twin chart columns with similar length, and the task completion time required to perform the chart inspection in the preliminary inspection phase. The experimental results showed that the subjects made the smallest number of repeated inspections of only 1.5 – 3. They resulted in the lowest error rate of about 1.2% when twin-column to non-twin column length ratio was less than 4:5 and minimum length of the column was 13 mm. In this case, the subjects needed only 4–8 s to accomplish the inspection task. The positive experience acquired within simulated gameplay conditions can be applied in developing novel didactic tools for training blind students in estimating linear dimensions of the virtual objects.

Acknowledgments

This work was financially supported by the Academy of Finland (grant 107278) and the project “Multimodal Collaboration Environment for Inclusion of Visually Impaired Children” funded by the EU Commission, IST-2003-511592.

References

1. Grabowski, N. A. and Barner, K. E.: Data Visualisation Methods for the Blind Using Force Feedback and Sonification. In: Part of the SPIE Conference on Telemanipulator and Telepresence Technologies V, Boston Massachusetts, (1998) 131-139.
2. Hatwell, Y., Streri, A., and Gentaz, E.: Touching for knowing. Cognitive psychology of haptic manual perception. Philadelphia, PA, USA: John Benjamins Publishing Company, 2003.
3. Ishii M. and Sato M.: A 3D Spatial Interface Device Using Tensed Strings. In: Presence-Teleoperators and Virtual Environments, MIT Press, Cambridge, MA, 3(1), (1994) 81-86.
4. Jia-Yu P., Hyung-Jeong Y., Faloutsos C., Duygulu P.: Automatic multimedia cross-modal correlation discovery. In: Proceedings of the 10th ACM SIGKDD Conference 2004, Seattle, WA, (2004) 653-658.
5. Lindemann R. and Tesar D.: Construction and Demonstration of a 9-String 6-DOF Force Reflecting Joystick for Telerobotics, In.: NASA International Conference on Space Telerobotics, 1989 (4) 55-63.
6. Phantom device. The information about the device may be found at: `http://www.sensable.com`.
7. Sribunruangrit N., Marque C., Lenay C., Hanneton S., Gapenne O., and Vanhoutte C.: Speed-accuracy tradeoff during performance of a tracking task without visual feedback. In: IEEE Transactions on Neural Systems and Rehabilitation Engineering, Vol. 12, No.1, (2004) 131-139.
8. Yu, W., Ramloll, R. and Brewster S.A. Haptic graphs for blind computer users. Haptic Human-Computer Interaction. Brewster, S.A. and Murray-Smith, R. (Eds.), Springer LNCS, Vol. 2058, (2001) 41-51.

"Assistec" – A University Course on Assistive Technologies

Kerstin Matausch, Barbara Hengstberger, and Klaus Miesenberger

University of Linz, Institute Integriert Studieren,
Altenbergerstrasse 69
4040 Linz, Austria
{kerstin.matausch, barbara.hengstberger,
klaus.miesenberger}@jku.at
www.integriert-studieren.jku.at

Abstract. Research results concerning Assistive Technologies show a growing demand of experts on AT deriving from an increase in use of Assistive Technologies which can be seen as an indirect result of worldwide population development trends. According to these recent changes efforts on inclusion of people with disabilities and older adults is of prime importance. Due to these facts the Institute Integriert Studieren started developing a new university course on Assistive Technologies. The characteristics are the composition of the course and its interdisciplinary content. Graduates will be awarded with an academic title. The following article describes the idea of the training, its contents, its realisation and its expected impacts.

1 Introduction

The last years are characterised by a constantly growth of awareness towards aging and disability. Demographic developments regarding Austria and other industrialised nations clearly show an increase of older adults and people with disabilities in general. A reason regarding growth of people with disabilities can be explained by the increasing danger of disability when aging [1].

Several economical sectors already reacted to these developments by starting to produce apparatuses especially for the older adult people following the ideas of Design for All.

Not only technical apparatuses for a convenient life are of interest for older adults, but also support systems to gain independence through life, as this group of population is more likely to be affected by a disability. Survey shows that older adult people aim autonomy in their lives [1]. People with disabilities and older adults are able to gain greater control over the own life by the use of Assistive Technologies (AT). They allow participation and more contribution to activities at home, school, work, leisure time or other communities as the Technology-Related Assistance for Individuals with Disabilities Act of 1988 declares [2].

Due to previous mentioned aspects, the use of Assistive Technologies will rise. To satisfy the growing demand, research shows that more and well educated experts on Assistive Technologies will be necessary.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 361 – 368, 2006.

Regarding the legislative framework there is no special law on Assistive Technologies in Austria, however an act on the equalization of people with disabilities. Austria avows itself to an equal treatment of people with and without disabilities. Equal treatment has to be realised in all spheres of life which means also a realisation in the field of Information and Communication Technologies (ICT). Assistive Technologies are an essential support to assure access to ICT, to equalization and additionally to improve employability.

All this motivated us to start working towards a new academic course and to close the gap between growing demand and offer in the German-speaking field of Assistive Technologies. We report on the academic course "Assistierende Technologien" - "Assistive Technologies". The course will include all necessary knowledge counsellors on Assistive Technology need and will comprise all categories of Assistive Technology as people working with AT presently often are experts for one category. For this purpose it contains technical knowledge which is necessary for employees in the areas of support, care and rehabilitation of people with specific needs and older adults. Therefore, graduates can be engaged in different areas of Assistive Technologies.

2 Current State

The current educational offer in the field of Assistive Technologies is inconsistent in Austria. Expertise is developed through both learning-by-doing and single seminars. Different enterprises working in the field of Assistive Technologies offer seminars, each lasting at maximum a couple of days. Another offering was SART [3], a summer academy on rehabilitation technologies, which took place in 1999. The institute Integriert Studieren already participated in its organisation and realisation.

Beyond Austria and the German speaking area there are for example the European wide network TELEMATE [4], online lectures in Italian language (SIVA) [5]. CSUN offers the so called ATACP, which is a detailed training [6].

Practical experience shows that AT counsellors need to know about products, adaptation, demand analysis, finances, funding and more. Derived from the present situation, a comprehensive education on Assistive Technology is meaningful.

3 The Course

The duration of the university course encompasses four terms. The first course will start in the winter term 2006 and the course language will be German. The course “Assistec” will be offered as an online eLearning application with a certain amount of mandatory attendance hours. The course can be referred as in-service training. Intentions for this kind of realisation are a high temporal and regional flexibility for participants, especially for employees. However, it supports lecturers too, for they save time and it was easier to involve experts from all over Austria. The course graduates will be awarded an academic degree called “Experts on Assistive Technologies”.

The university course aims to educate people deriving from different vocational backgrounds regarding the special field of AT. Graduates will be experts in the area of Assistive Technologies especially concerning assortment of appropriate AT, usability of

AT, funding, application, adaptation, management and service and counselling respectively. Moreover, the course stresses concentrated and goal-oriented transfer of knowledge according to the up-to-date state of the art in a multidisciplinary environment. One major key feature and goal as well is that the course substantially emphasises practical training and application of the theoretically gained knowledge. Moreover, one major intention is to enhance quality in the practical treatment regarding the profession fields of health care and support and services of people with disabilities. In addition to that the implementation of the university course also aims to improve and foster product development of Assistive Technologies.

The course is intentionally appealing to people from different vocational and educational backgrounds. Therefore the target groups addressed by the university course are multifaceted as follows:

- *Vocational field of welfare*
  Within this vocational field the course is especially addressing people employed in the field of "people with disabilities" and "integration of people with disabilities" dealing with counselling, care, support, service and accompaniment of people with disabilities.
- *Vocational field of health care*
  People who are working in the fields of rehabilitation, nursing, care and support of people with disabilities and older adult people are approached in particular.
- *Vocational field of education*
  Within this target group we are appealing to both teachers of standard schools and adult education as well as special school teachers and pedagogues dealing with children with disabilities.
- *Vocational field of Assistive Technologies*
  Especially addressed are people who are engaged in the areas of production, distribution and trading, maintenance, training, research and development of Assistive Technologies.

To be in accordance with the idea of an equal access to education offers, the Institute Integriert Studieren encourages in particular people with disabilities to participate in the university course "Assistec". One more intention is to enhance vocational chances of people with disabilities at the open labour market due to an upgraded qualification.

3.1 Curriculum

The university course's curriculum consists of four modules. Each of the modules is composed of single seminars. As a whole the university course comprises 18 seminars. The following figure outlines the university course's contents:

Module one is focusing on imparting fundamental knowledge concerning medicine, physiology and classification of disability, legal foundations regarding disability and funding facilities in Austria, Assistive Technologies and finally Design for All. As the module's name already implies, the contents constitute fundamentals for the course at whole and especially for the following subject.

Contents of "Assistec"

module 1: Fundamentals
eLearning
Medical and Physiological Fundamentals
Legislative Framework and Funding
Design for All
Fundamentals AT & Reha-Technology

module 2: Assistive Technologies Special Knowledge
Aids for Specific Learning Difficulties
Aids for Hearing Impaired
Aids for Vision Impaired
Mobility Aids
Augmentative or Alternative Communication
ATs in different areas (work, home, education, leisure time)
Practical Experience

module 3: Process of Assortment and Provision of AT & ICT
Assessment and Demand Analysis
Environment Analysis in Technical Area, Sociological Area, Economical Area
AT-Management and Mediation

module 4: Assistive Technologies in Practice and Application
Practical Course and Thesis
Research in the Field of AT and Future Developments

Fig. 1. Contents of the university course "Assistec". The mentioned components accord to the contents and do not exactly reflect the whole seminar titles.

Teaching contents of module two are emphasising on special knowledge of Assistive Technologies including practical training units regarding AT products and their application. Therefore, this module disposes of a central position in the whole curriculum.

Module three is dealing with the management and realisation of the process of assortment and provision of Assistive Technologies, as a goal is to educate counsellors and process managers. Hence pivotal issues are needs assessment, analysis of the environment of people with disabilities in respect of technical, sociological and economical areas and mediation.

During module four the participants have to undergo practical experience by managing and documenting a process of assortment and provision of AT by means of composing a scientific thesis. This kind of work placement also can be completed among holding down a job as the whole university course is organised as an in-service training.

A special feature of the developed curriculum is the interdisciplinary composition of professional subjects and contents of teaching respectively. The range of subjects highlighted covers medical, legal, technical, economical and management aspects, sociological aspects as well as psychological aspects. One more crucial element of the course is the cooperation and networking with enterprises in the field of AT.

3.2 eLearning System

The design of the course is characterised by a blended learning system, which "combines face-to-face instruction with computer-mediated instruction" [7].

According to this definition blended learning is used regarding the university course "Assistive Technologies" as a combination of online learning and presence learning elements.

As the university course should be open for all people equally regardless of a possible disability the issue of accessibility has to be raised. As a fully accessible system is not to be found at the market, the idea occurred to adapt an already existing eLearning system. An open source course management system (CMS) called "Moodle" [8] has been chosen due to its good results according to a first accessibility evaluation following the Web Content Accessibility Guidelines [9] published by the Web Accessibility Initiative [10].

The used eLearning System has been evaluated several times by people with specific needs. Nevertheless the fact is that the evaluation and the gathered experiences using and testing this system still show deficits concerning full accessibility. Great attempts are being undertaken to achieve a fully accessible version until the start of the course.

The eLearning system acts both as a communication platform and a platform where all study materials are available in accessible formats. Concerning the functionality "Moodle" contains a wide range of features:

- Assignment Module:
  Teachers can give homework assignments and students can upload their assignments in any file format to the server.
- Chat Module:
  The chat module allows synchronous text interaction.
- Forum Module:
  Different types of forums are available for students and teachers.
- Quiz Module:
  Teachers can define a database of questions for the students.
- Resource Module:
  Teachers can provide scripts, slides, lecture notes, videos etc. supporting their lecture.
- Survey Module:
  Students can evaluate single courses and feedback on the students' results is provided.

Even though the eLearning system "Moodle" plays a crucial role within the blended learning settlement, the phases of personal attendance also are of great significance. Mandatory presence is planned for averaged three times per term with duration of two days. The phases of presence are on the one hand used in order to initiate social contacts and formation of groups and on the other hand to demonstrate and present practical and theoretical contents.

Overall, the blended learning settlement is enabling a high level of regional and temporal flexibility for the participants – and this is especially beneficial for people with disabilities in respect of access to the course itself and mobility issues.

3. 3 Graduates

The qualification profile of the graduates is characterised by the following features:

- Graduates gain comprehensive and scientific specialist knowledge in the area of Assistive Technologies and their application
- Graduates acquire a verifiable degree
- Graduates have the ability to provide people with disabilities and older adults with adequate Assistive Technologies and to support them through the process of assortment and provision of AT
- Graduates are qualified to self-contained organisation, coordination, management and handling of the whole process of assortment and provision of Assistive Technologies
- Graduates are opening up of a new field of profession due to their interdisciplinary knowledge concerning the social, technical, medical and rehabilitation areas
- Graduates are highly aware of the societal context of disability, aging and Assistive Technology
- Graduates own knowledge regarding the legislative framework and funding possibilities in respect of the issue of disability associated with Assistive Technology
- Graduates have an increased sensibility and social competence in interacting and dealing with people using Assistive Technologies

Keeping in mind this profile of qualifications the graduates of “Assistec” gain during their education the vocational field of “Experts on Assistive Technologies” can be specified by some criteria. First of all, they are the one central contact person for clients and users of Assistive Technologies. Having one main contact person arranging all aspects of the assortment and provision of AT is a crucial advantage for people provided with AT. The experts are independent counsellors who are not sales oriented but do have a detailed overview on the whole range of AT products and choose the most adequate device for the clients. Furthermore, AT experts are process managers. That is to say that they are empowered to organise and coordinate the provision process of AT taking into account juridical, medical, technical, economical and sociological aspects. Due to this they have leadership and management skills. Moreover they are a representative of the user group of AT as well as of economy referring to AT organisations producing and distributing AT. In addition to that these experts act as multipliers in their vocational field in the respect as they fulfil the task of awareness raising and sensitisation. Finally a significant aspect of the expert’s vocational field is the usage of mediation and conflict management skills if conflicts and difficulties emerge during the process of assortment and provision of AT.

The chart below shows the influence and impact of the developed university course “Assistec” and its graduates on wider society.

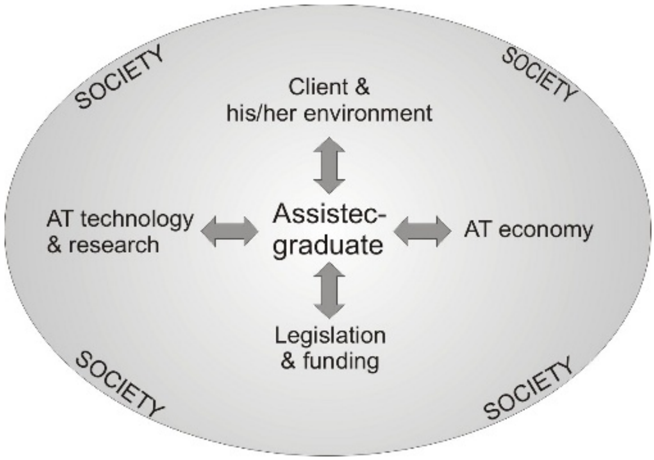


Fig. 2. The figure shows a graduate's connection and interexchange with various target groups such as the client, economy and research in the field of AT as well as legislation and funding

4 Impact

This course will be a contribution of the Institute Integriert Studieren at the University of Linz towards current efforts on inclusion of people with specific needs and population development trends. We expect to change the present vocational field by offering a new special training. Graduates have competencies to manage a process of assortment and provision from the very beginning till required trainings and maintenance relating to all categories of Assistive Technologies. We also expect to improve quality regarding care of older adults and people with specific needs as well as giving impulse to product development due to an intensive communication between clients and "Academic Experts on Assistive Technologies".

Graduates possess a high sensibility and social aptitudes in interaction with people who use Assistive Technology, as they concentrate on disability, interact with concerned and know about the effects of technologies on self-determined life. For this, they are aware of disability's societal effects and Assistive Technology's context.

Graduates push awareness and sensibility towards Assistive Technologies in their environment because of their contacts to clients. On account of this, they will be multipliers in their vocational fields.

Acknowledgement

The development of the university course on Assistive Technologies has been funded by the European Social Fund and bm:bwk.

References

1. Lifetool: Informations- und Kommunikationstechnologie für Menschen im Alter, http://www.lifetool.at/rte/upload/6_Fachforum/IKT_studie_2004_Endbericht.pdf
2. Technology-related Assistance of Individuals with Disabilites Act of 1988 as amended in 1994 http://www.resna.org/taproject/library/laws/techact94.htm
3. SART - Sommerakademie für Rehabilitationstechnik http://www.fortec.tuwien.ac.at/reha.d/projects/sart/sart.html
4. TELEMATE http://www.telemate.org/
5. SIVA http://www.siva.it/eng/education/default.htm
6. CSUN http://www.csun.edu/codtraining/index.htm
7. Graham, C.R.: Blended Learning Systems. In: Bonk, C. J., Graham, C. R. (eds.): Handbook of blended learning: Global Perspectives, local designs. Pfeiffer Publishing, San Francisco, http://www.publicationshare.com/graham_intro.pdf
8. Moodle - an open source eLearning-System, www.moodle.org
9. Chisholm, W., Vanderheiden, G., Jacobs, I. (eds.): WCAG 1.0 - Web Content Accessibility Guidelines 1.0: http://www.w3.org/TR/WAI-WEBCONTENT/
10. WAI - Web Accessibility Initiative: http://www.w3.org/WAI/

People with Disabilities: Entertainment Software Accessibility

Introduction to the Special Thematic Session

Dominique Archambault

Université Pierre et Marie Curie, Inserm U483/INOVA,
9 quai Saint Bernard, 75252 Paris cedex 5, France

The commercial market for computer games and other multimedia products is extremely large and young people have a considerable experience of such games. Disabled users have very limited access to this important part of the youth culture. Indeed there are few entertaining computer games which are accessible for them. Research and development in the field of IT and the disabled has focused on education rather than leisure.

During the last five years a few projects have been exploring various ways of rendering game situations for users with special needs. Specific games or specific versions of mainstream games were developed, evaluated and for some of them distributed.

The TiM project [1], funded by the European Commission, created a game development API [2], "blindstation", which helps to design accessible games, together with a set of games which were used to study various game situations with users. One of these games, *Mudsplat* is currently available for download:

`http://www.timgames.org/mudsplat`

One can find a good number of audio games. The Swedish Library of Talking Books and Braille (TPB) has published web-based games dedicated to young children with visual impairment [3]. On the other end, Terraformers is the result of three years of practical research in developing a real-time 3D graphic game accessible for blind and low vision gamers as well as full sighted gamers [4]. A quite comprehensive list of audiogames can be found at:

`http://www.audiogames.net`

Various kind of tactile and haptic devices have been explored too [5,6,7]. [8] proposes a game generator to create simple audio/tactile games dedicated to very young children. [9] reports experiences about virtual reality games involving blind users carrying devices in a backpack and moving on a real football field.

In the paper entitled *"Semi automatic generator of tactile video games"*, Alexis Sepchat will present his work about the representation of a 2 dimensional game space on a linear Braille display, and his models to navigate and play using this kind of devices.

From the experiences collected, we can bring together a set of general rules allowing to improve the accessibility of mainstream games [10,11]. The IGDA (International Game Developer Association) published a white paper about accessibility of mainstream games [12].

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 369–371, 2006.

Mathew Atkinson's paper, *"Making the Mainstream Accessible: What's in a Game?"*, based on his experience in AudioQuake (first-person shooter games accessible to visually impaired people), discusses some of the low-level accessibility infrastructure employed in this game and compares it to other contemporary research.

Dimitris Grammenos, in the paper *"Access Invaders: Developing a Universally Accessible Action Game"*, discusses the notion of Universally Accessible Games. The paper is based on a case study (the Space Invaders game).

In the paper *"Internet and accessible entertainment"*, Morten Tollefsen will present the development of an Internet based game, *"HeiPipLerke"*, which is used as an example of methodological approach in the framework of the UPS project [13].

The next step is to write guidelines for accessibility of Games, and to make them accepted and used by mainstream game developers. These guidelines will allow:

- To facilitate the development of specific software for disabled users
- Entertainment software publishers to develop new products that respect as well the reality of the market and the needs of all users. Indeed young people with disabilities wish to use the popular computer games used by friends and family.

During the *Workshop on Accessible Games* held in Linz in November 2005, this task was set up as an international collaboration, leaded by University of Linz and MediaLT. A working group was created and the framework of the accessibility guidelines was outlined. In the paper *"Guidelines for the Development of Accessible Computer Games"*, Roland Ossmann will report the progresses of the working group. The current state of the guidelines will be presented, together with the system used to work on these guidelines.

References

1. Archambault, D.: The TiM Project: Overview of Results. In Klaus, J., Miesenberger, K., Zagler, W., Burger, D., eds.: Proc. ICCHP 2004 (9^{th} International Conference on Computers Helping People with Special Needs). Volume 3118 of Lecture Notes in Computer Science., Springer (2004) 248 – 256
2. Sablé, S., Archambault, D.: Blindstation: a Game platform adapted to visually impaired children. In Craddock, G., McCormack, L., Reilly, R., Knops, H., eds.: Assistive Technology – Shaping the future, Proceedings of the AAATE'03 Conference, Dublin, Ireland, Amsterdam, The Netherlands, IOS Press (2003) 232 – 236
3. Eriksson, Y., Gärdenfors, D.: Computer games for children with visual impairments. In: Proceedings of the Fifth International Conference on Disability, Virtual Reality and Associated Technologies, Oxford, UK (2004) 79 – 86
4. Westin, T.: Game accessibility case study: Terraformers - a real-time 3d graphic game. In: Proceedings of the Fifth International Conference on Disability, Virtual Reality and Associated Technologies, Oxford, UK (2004) 95 – 100

5. Sjöström, C.: The sense of touch provides new computer interaction techniques for disabled people. Technology and Disability (IOS Press) **10**(1) (1999)
6. Raisamo, R.: Haptics research in tauchi: Aiding visually impaired children and enabling haptic interaction. In: Hands on Haptics: Exploring Non-Visual Visualisation Using the Sense of Touch, a CHI 2005 workshop, Partland, Oregon (2005)
7. Evreinov, G., Evreinova, T., Raisamo, R.: Mobile games for training tactile perception. In Rauterberg, M., ed.: Entertainment Computing - ICEC 2004, Third International Conference, Eindhoven, The Netherlands. Volume 3166 of Lecture Notes in Computer Science., Springer (2004) 468 – 475
8. Puret, A., Archambault, D., Monmarché, N., Slimane, M.: A simple game generator for creating audio/tactiles games. In Pruski, A., Knops, H., eds.: Assistive Technology: From Virtuality to Reality, Proceedings of the AAATE'05 Conference, Lille, France, Amsterdam, The Netherlands, IOS Press (2005) 545 – 549
9. Velleman, E., van Tol, R., Huiberts, S., Verwey, H.: 3d shooting games, multimodal games, sound games and more working examples of the future of games for the blind. In Klaus, J., Miesenberger, K., Zagler, W., Burger, D., eds.: Proc. ICCHP 2004 (9^{th} International Conference on Computers Helping People with Special Needs). Volume 3118 of Lecture Notes in Computer Science., Springer (2004) 257–263
10. Archambault, D., Olivier, D., Svensson, H.: Computer games that work for visually impaired children. In Stephanidis, C., ed.: Proceedings of HCI International 2005 Conference, Las Vegas, Nevada (2005) 8 pages (proceedings on CD-Rom).
11. Archambault, D., Olivier, D.: How to make games for visually impaired children. In: Proceedings of ACE05 Conference (ACM SIGCHI International Conference on Advances in Computer Entertainment Technology), Valencia, Spain (2005) 450 – 453
12. International Game Developers Association (IGDA): Accessibility in Games: Motivations and Approaches. (2004)
13. Tollefsen, M., Lunde, M.: Entertaining Software for Young Persons with Disabilities. In Klaus, J., Miesenberger, K., Zagler, W., Burger, D., eds.: Proc. ICCHP 2004 (9^{th} International Conference on Computers Helping People with Special Needs). Volume 3118 of Lecture Notes in Computer Science., Springer (2004) 240 – 247

Semi Automatic Generator of Tactile Video Games for Visually Impaired Children

Alexis Sepchat[1], Nicolas Monmarché[1], Mohamed Slimane[1], and Dominique Archambault[2]

[1] Laboratoire d'informatique de l'Université Francois Rabelais de Tours, France
alexis.sepchat@univ-tours.fr, nicolas.monmarche@univ-tours.fr, mohamed.slimane@univ-tours.fr

[2] Université Pierre et Marie Curie, INSERM U483\INOVA , France
dominique.archambault@upmc.fr

Abstract. Currently, few video games are accessible for visually impaired people. Nevertheless, there are two ways in order to improve video games accessibility : the use of sound or the use of touch. Even if the latter turns out to be the main exploited solution, the use of touch remains substantial. Indeed, touch is the base of Braille learning and Braille knowledge is the only way for visually impaired persons to access written information alone without any technical help such as computer and vocal analysis [1]. This article introduces our works about tactile video games. It shows games like Snake or Maze, which can be played from a Braille display. Finally, these works have led us to think about the way to introduce tactile games as play aspect in Braille learning[2,3].

1 Introduction

Created during the first half of the 19th century by Louis Braille, Braille was an important designing issue for visually impaired people. Nevertheless, the volume of transcribed data is significant. That is the reason why the association of computer science and Braille terminal (in the beginning of the 80's) has been a great revolution in the development of Braille. It allows the association of both an accessible representation technique and a storage one.

But, if at its beginning, computer science was easily accessible because of the simplicity of its interfaces. Its improvement has led to a loss of the numeric accessibility. So, at present, if it is a wonderful learning tool for visually impaired people, it above all remains a working tool and does not have an entertaining aspect as for other people without any visual deficiencies or with other kinds of deficiencies. Yet, computer games are a wonderful tool to learn [2].

In previous works in the bosom of the Computer Science Laboratory (L.I.) of the University of Tours, we have already focused on this entertaining aspect. We have developed games based on sounds and textures for young children and the corresponding game generator [4]. These works were a part of the TiM project (Tactile Interactive Multimedia) [5] which goal is the development of accessible video games for blind and partially sightless children.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 372–379, 2006.

This new study focuses on games based on tactile modality and more especially the use of Braille terminals [6]. It means that the average age of the target audience is higher than in our previous study. The development of such games has led us to the conception of a semi automatic generator of such games.

This article is organized as follows: in section 2, we will introduce some generalities about accessible video games. Then in section 3 and 4, we will give details about the semi automatic generator of tactile accessible video games for visually impaired children and the remaining works. In section 5 we will show some examples of games and representations developed from this semi automatic generator. In section 6, we will introduce few results of our games tests. The last section will be devoted to the area of improvement that we have identified and conclude this paper.

2 Accessible Video Games

Standard video games are not, or at least not easily, accessible. These difficulties are linked to players' perception problems but also his/her abilities (mental or motor): speed or complexity of the game. Our works, based on visual impairment, focus essentially on the perception problems.

Visual impairment prevents player from using standard output devices. So the main problem of visually impaired player is to acquire the game information. Two modalities can be used to replace visual modality : tactile modality or sound modality. Till now, sound one keeps being the most important modality, almost all accessible video games or projects about accessible video games use it[1]. Nevertheless, tactile modality is very important. Not only could its use, thanks to Braille terminals (Cf. figure 1), permit to improve the touch sensibility of the visually impaired players but it could also be used as a entertaining aspect in the learning of Braille.

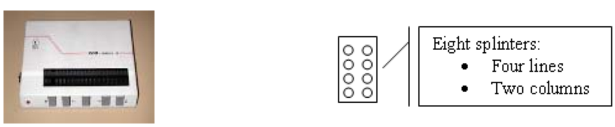


Fig. 1. Braille terminal : ECO ONCE 20 - Braille cell details

Moreover, such equipment (Braille terminals) is very expensive. Consequently, people can not get several ones and the created games have to run whatever the Braille terminal used. We have thus decided to use the Libbraille library [7].

Besides, a second very important aspect in accessible video games is to take several interfaces into account. Indeed, everybody must be able to play with these video games and not only visually impaired persons. So accessible video games developers must not create a new video games universe but a sharable universe

[1] http://www.audiogames.net/

in which impaired persons and others could play together. Several projects take this aspect into account [8,9]. A way to create such a universe is to separate the game itself (data, structure, ...) and its interfaces (input and output interfaces). This separation would facilitate the development of new interfaces, taking into account particular impairments or not , and following several guidelines[2,3,4].

3 The Semi Automatic Generator

3.1 Introduction

Although most of the accessible video games for visually impaired children are based on sounds, the use of the touch modality is very important, in particular as play medium as part of Braille learning. But several difficulties are encountered during the development of tactile video games using a Braille Terminal. These difficulties are linked with the transcription of the game on the Braille terminal:

- the game dimension and the Braille terminal dimension are different. Indeed, games chosen are 2D ones whereas the Braille line can be considered as 1D environment or as a very restricted 2D environment;
- the limited number of Braille cells. All the game can not be transcribed on the Braille line. So the most important pieces of information have to be extracted;
- the structure of the Braille representation of the game must be easy to read because the player needs to read it quickly.

In order to facilitate the development of new games, as well as, the creation of new Braille representations for an existing game, we have decided to develop an automatic generator of such games. Its aim is to develop new games without programming their main structure, just focusing on the most important points like Braille representations. So everyone, with just a few knowledge in computer science, can create its own games and especially its own representations: medical personal, educators, parents, ...

Obviously, the developed games must have a similar structure (internal structure or visual interface) and all the games have to follow the same philosophy. The one we have decided to use is the movement of a character inside a grid. A grid can be considered as a whole of sectors where each sector is in a given state. The different states are limited and they depend on the game. Even if this choice can be considered as very restrictive in the development of next games, finally most of them can be perceived like this.

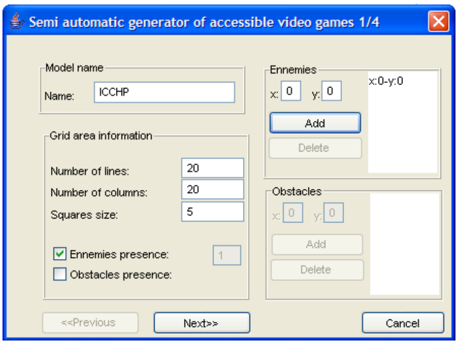
For example, we have already developed two games (and several Braille representations with each one): a Maze Game and a Snake Game (Cf. section 5). But most of games can be considered like these ones: sport ones, platform ones,...

[2] (Roland Ossmann) - http://gameaccess.medialt.no/guide.php
[3] (Medialt) - http://www.medialt.no/rapport/entertainment_guidelines/index.htm
[4] (IGDA) - http://www.igda.org/accessibility/IGDA_Accessibility_WhitePaper.pdf

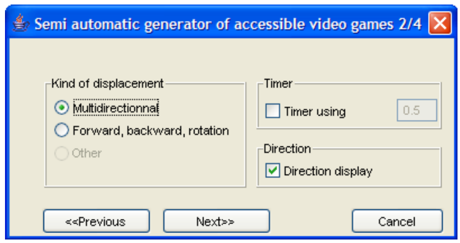
Unfortunately, in the current version, this tool is still a semi automatic generator. Therefore, the result is reduced to a skeleton of the expected game. It remains necessary to manually finalize this game. But the work is amply facilitated as we will demonstrate in the following section.

3.2 The Several Forms

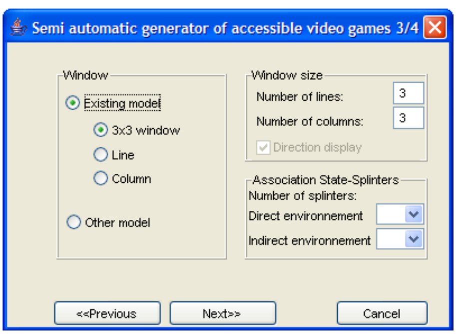
This semi automatic generator is based on several forms (4) allowing the access to information about the game to build.



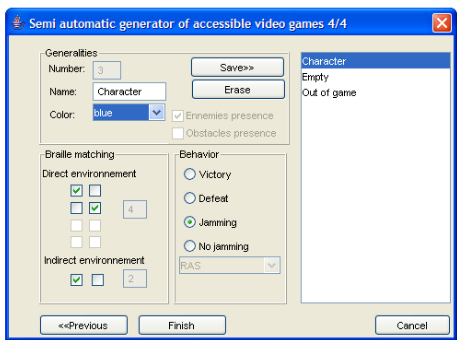
Semi automatic generator: Step1



Semi automatic generator: Step2



Semi automatic generator: Step3



Semi automatic generator: Step4

First Step : General Information. The first form deals with general information about the game like its name but it essentially focuses on the grid of the game.

Indeed, one game is associated with one grid, which means that each game has to define its own grid i.e. its size (i.e. its size - number of sectors per line and per column) and the state of each sector of the grid. Nevertheless, a project of automatic generation of grids is in progress.

Then, the square size is linked with the representation of the game on the visual interface: each sector is represented by a square in the visual interface.

This information allows to define its size. It can be interesting to regulate this size to give partially sighted children the possibility to play these games.

Second Step : Moving. This second form deals with the moving of the character inside the grid.

This form permits to choose the character way of moving when the player hits a multi-directional control key or a timer expired. Two ways of moving are available:

– Multi-directional: The character moves according to the control key hit.
– Forward(key "up"),Backward(key "down"),Rotation(keys "left" and "right" : modify the current character direction in clockwise or anticlockwise): The character moves according to its current direction and the control key hit (forward if a timer expired). In that case, it is necessary to transcribe the direction so this information will appear in the first cell of the Braille representation.

Third Step : Window and Braille Cell Structure. A Braille terminal has a limited number of Braille cells. So, the whole game can not be transcribed into the Braille representation and some pieces of information have to be selected.

On the one hand, this selection begins with the definition of a window containing all the sectors which have to be transcribed (these sectors are called "direct environment"). This window is centered on the character and can take many patterns like a line, a column (the current line or the current column of the character), a 3x3 window, ...

Fig. 2. Several patterns of window

Moreover, the number of sectors contained in this window depends on the number of splinters associated with each sector. This number of splinters depends itself on the number of states available in the game because each state has to get one only Braille representation. Finally, the more states there are, the more splinters are necessary and the less sectors can be transcribed.

For example, in a game with 9 states, it is possible to transcribe these 9 states with 4 splinters ($2^4 = 16$ and $16 > 9$). So, only 2 sectors could be transcribed in the 8 splinters of a Braille cell ($2 \times 4 = 8$). Whereas in a game with only 4 states, they can be transcribed with 2 splinters ($2^2 = 4$) so 4 sectors could be transcribed in a Braille cell ($4 \times 2 = 8$).

On the other hand, it can be interesting to keep some splinters free in order to transcribed information (i.e. sectors) not in the initial window. Moreover, if this information is not in the initial window, it is supposed to be far enough from the character to accept a loss of precision. So it is possible to cluster states into classes. Finally, the number of classes is less important than the number of states. So, it is possible to associate new configurations of splinters containing less splinters to each class.

For example, in a game with 9 states, 4 splinters are associated with each state. If these 9 states are clustered into 4 classes. Each class are associated with 2 splinters. Finally, each Braille cell transcribes one sector from the direct environment (contained in the initial window) and two sectors from the indirect environment ($1 \times 4 + 2 \times 2 = 8$).

Finally, this form allows to define the pattern of the initial window and the number of splinters associated with each environment (direct and indirect one).

Fourth Step : States Management. This last form allows to control each state of the game : its name, its color, the corresponding configuration of splinters or the behavior of the game when the character encounters this state during its movement.

4 Games Skeleton

Finally, this semi automatic generator builds a Python skeleton of the game wished. The programmer (or the person with a few notions of programming) needs to finalize the game. The main points of this finalization consist in:

– placing the several elements on the grid (currently just the obstacles and the enemies can be place from the generator);
– configuring the behavior of the character during its movement (when it hits the different elements i.e. states);
– configuring the building of the Braille representation i.e. the way of browsing the initial window and the location of each information on the Braille cell.

5 Snake and Maze Games

Two tactile video games based on the moving of a character inside a grid and several representations have been developed with this tool:

1. *a Maze game:* the player has to find the exit, collecting items and avoiding enemies and obstacles;
2. *a Snake game:* the player is a snake and has to find and eat apples without hitting its tail or wall. When the snake eats an apple the score is up but the length of its tail increases too.

To illustrate the building of a Braille representation, the following example has been extracted from the Maze game: In that game 7 states are available (character, empty, out of game, obstacle, enemy, item and exit) and have been associated with 4-splinters configurations ($2^4 > 7$ - Cf figure 3). Then, these 7 states have been clustered into 4 classes (Empty, Jamming : obstacle or out of game, Interesting : exit or item, Danger : enemy) each one represented with 2-splinters configurations ($2^2 = 4$ - Cf figure 3).

In that example, the initial window is the current character line (Cf figure 4).

Finally, a Braille representation example could be the following (Cf figure 5).

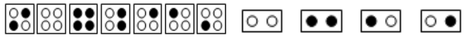


Fig. 3. Maze Game example: 4 and 2-splinters configurations

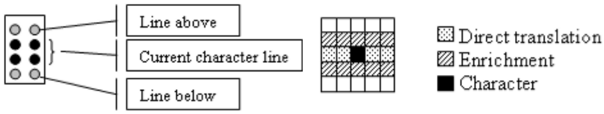


Fig. 4. Maze Game example: Braille cell structure and corresponding window

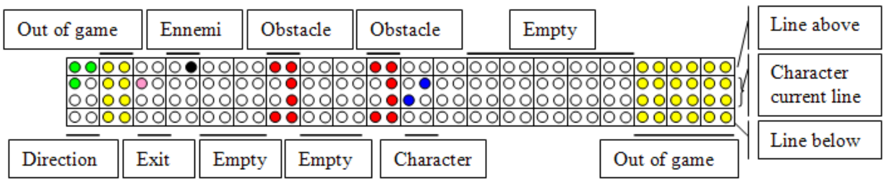


Fig. 5. Maze Game Braille line example

6 Evaluation

In order to validate our games, few tests have already been realized with several visually impaired adults who are advanced Braille reader but also with people without any visual impairment (playing just from the Braille representation of the graphic interface as in figure 5).

On one hand, people without any visual impairment understood very quickly the game because we have used the whole game grid interface simultaneously with the Braille representation one to explain them the functioning of the game. They have met no difficulty in understanding the Braille representation building.

On the other hand, this explanation was more difficult with visually impaired player. But after few minutes of explanations (between 5 and 10 minutes), most of them played without any difficulty and several ones even suggest new Braille representations. Few players did not manage to play because of a lack of time for explanations or because of additional motor impairment.

7 Conclusion

This tool is a semi automatic generator, a programmer's intervention remains necessary. We have made this choice in order to develop a friendly-use tool to use but especially to keep a large range of personalization. Indeed, a full generator would be more complex to use because it would have to take into consideration many particular cases in order to create full games. Consequently, forms would have to be more complete. Moreover, more complete forms would reduce the generation of new games because it would prevent the programmer from creating new scenarios.

Additionally, these games are still exclusively based on the tactile modality. It would be interesting to develop games using simultaneously tactile and sound modalities. This combination would permit to develop new representations offering a new element to transcribe information.

This study proves that is possible to create a new kind of games based on Braille displays. Even if Braille representations have to be simplified to focus on Braille reader beginner or easier games have to be used (as game based on words, . . .), it would be very interesting to insert them as entertaining aspect into Braille learning.

To conclude, the tactile modality is substantial in visual impairment universe but there are currently very few video games using this modality. Our works have already led to the development of two tactile video games, but especially to the development of a semi automatic generator. The aim of this tool is to facilitate the creation of such games and to help visually impaired children to develop their touch, to familiarize themselves with Braille terminals and to help them in Braille learning through games.

References

1. CNIB: The importance of braille literacy. (The Canadian National Institute for the Blind)
2. Mitchell, A., Savill-Smith, C.: The use of computer and video games for learning. The Learning and Skills Developpment Agency, London (2001)
3. The State Of Queensland (Department of Education): Teacher aides working with students with disabilities. Book G: Vision impairment. The State Of Queensland (Department of Education) (1998)
4. Puret, A., Archambault, D., Monmarché, N., Slimane, M.: A simple game generator for creating audio/tactile games. In: Assistive Technology: From Virtuality to Reality (AAATE'2005), Lille, France (2005) 545–549
5. Archambault, D.: The TiM Project: Overview of Results. In: 9th International Conference on Computers Helping People with Special Needs (ICCHP'2004), Paris, France (2004) 248–256
6. Sepchat, A., Monmarché, N., Archambault, D., Mohamed., S.: Accessible video games for visually impaired children. In: The Third Annual International Conference in Computer Game Design and Technology Workshop (GDTW'2005), Liverpool, UK (2005) 58–67
7. Sablé, S., Archambault, D.: Libbraille: a portable library to easily access braille displays. In: 8th International Conference on Computers Helping People with Special Needs (ICCHP'2002), Linz, Austria (2002)
8. Westin, T.: Game accessibility case study: Terraformers a real-time 3D graphic game. In: 5th Intl Conf. Disability, Virtual Reality And Assoc. Tech. (ICDVRAT'2004), Oxford, UK (2004) 95–100
9. Grammenos, D., Savidis, A., Stephanidis, C.: UA-chess: A universally accessible board game. In: HCI: Exploring New Interaction Environments - Volume 7 of the Proceedings of the 11th International Conference on Human-Computer Interaction (HCI International 2005), Las Vegas, Nevada, USA (2005)

Making the Mainstream Accessible: What's in a Game?

Matthew T. Atkinson[1], Sabahattin Gucukoglu[2], Colin H. C. Machin[1], and Adrian E. Lawrence[1]

[1] Loughborough University, Loughborough, Leicestershire, LE11 3TU, England
{M.T.Atkinson, C.H.C.Machin, A.E.Lawrence}@lboro.ac.uk
[2] The AGRIP Project
sebby@agrip.org.uk
http://www.agrip.org.uk/

Abstract. Though accessible gaming is a well-established phenomenon, few mainstream applications of it exist. We present some of the work of the AGRIP project – an effort to develop techniques to render modern first-person shooter games accessible to the blind and vision-impaired. We discuss some of the low-level accessibility infrastructure employed in the game AudioQuake and compare it to other contemporary research. The project's ultimate goals of generalisation and use of the technology in educational settings are also introduced.

1 Introduction

AudioQuake is the first adaption of an existing mainstream game designed specifically for sighted people that has been made playable by blind gamers. It is unique in terms of the range of Internet-enabled gameplay modes it provides. At one level, it could be termed an "accessibility layer" for Quake[1].

This paper describes the work of the AGRIP project – an effort to develop techniques for making mainstream games, tools and their communities accessible to blind and vision-impaired gamers that has been active since May 2003. The approach taken by this project contrasts with other contemporary research [1,2,3] in the following ways.

- *Adaption* - Whilst other projects often develop engine and game platforms anew, this project modifies existing well-designed mainstream technology to improve its accessibility and usability for all. The goal is *not* to retrofit accessibility, but to show how properly-designed systems may be made accessible through the processing and rendering of information at separate levels.
- *Generalisation* - An important aim is to use the project to develop a deeper understanding of accessibility barriers that can be used in the development of general techniques to deal with those issues. We hope to create "portable" solutions to accessibility problems, that can be used in other (academic and leisure) settings.

[1] The seminal first-person shooter from id Software.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 380–387, 2006.

Throughout development, the AGRIP project has been shaped by community feedback. This incorporates that given by users via e-mail and using project mailing lists. Many comments from users and suggestions for improvements have been made. Using the mailing list approach has enabled a number of interesting discussions to take place *between* users of the software, thus giving us a greater insight into how effective the work has been. AudioQuake has been used as the basis of a number of educational workshops at the 2005 International Computer Camp for Vision-Impaired People.

The rest of this paper expands on the user requirements, design, technical and other issues encountered in providing low-level access to the game (essentially local navigation).

2 Local Navigation

There are two main strands of navigation present in almost any computer game, and in any problem that involves searching a solution space. These are *global* navigation towards one's ultimate goal – the silver key; the red team's flag; the shortest path between nodes in a tree – and *local* navigation – "How do I get out of this room?". At the low-level game accessibility stage of the project's development, the primary concern was to develop the mechanisms to support effective local navigation.

In this style of game, the player controls their character directly, and is presented with the character's world view. A number of "devices" which act as navigation aids are provided to facilitate accessibility. These devices attempt to resemble real-world mobility aids where possible and encourage the user to navigate the virtual world in a similar way that they do in the real world. For example, to allow players to get a very fine-grained sense of what obstacles/structures surround them, a "sweep" can be performed. This gives similar feedback to that which a mobility cane might (albeit in sound), though with a longer range.

This method of increasing accessibility has enabled users of different abilities to use only the devices they need. This approach can be found in some other fast-paced accessible games such as GMA's "Shades of Doom".

3 Structure Adaption and Filtering

Many existing accessibility systems are tightly coupled to an underlying mainstream technology used by non-disabled people. They provide accessibility by processing the output created by a given mainstream system and adapt it for users with certain disabilities (this is the way that screenreaders and accessible PDF viewers work, for example). This can cause a number of problems such as:

- Much work is required to interpret the meaning of the original output, based on visual markup such as layout and colouring [4,5].
- It may not be possible to directly extract or infer the types of information that specialist/disabled users require from output targeted at the "normal" user.

– When the underlying technology becomes obsolete, so does the accessibility system built on top of it[2].

The examples above fit in with many contemporary approaches, which emphasise the idea that to truly cater for the varying needs of users, the information behind any final output must be adapted for those needs. Rendering should be decoupled from information processing so that "accessibility" systems can utilise data at the same level that mainstream systems currently do.

3.1 Domain-Specific Solutions

In the case of the current work, this was achieved by utilising the architecture of modern computer games. Principles for *how* the information must be adapted were developed. Some domain-specific principles of structure simplification we have developed are described below. It would be very useful if the principles could be made more general; work towards achieving this is being carried out.

– *Necessity-Based Rendering* - It is not required, or even useful, to bombard the user with an audio interpretation of all graphical output from the game. In the context of navigation, obstacles such as walls need only be explicitly rendered[3] when they *become obstacles* for the user. Doors or ramps, however, should always be rendered, even if they are not directly in front of the player.
– *User-Centred Flexibility* - Even when considering only blind gamers, we found significant diversity in terms of preferences for rendering style. Two main schools of thought existed[4]. Within these, users had differing opinions on factors such as how quickly object indicator sounds should fall off with distance, how often scans for objects should be carried out and so on.

Note that through adhering to these principles, no information has been explicitly *added* – this does not appear to be necessary to help disabled people overcome barriers to accessibility. Our and others' work [6,7] has shown that simply adapting the underlying information and transforming it into the most appropriate format for rendering is the most important factor in increasing accessibility.

In fact, feedback given by AudioQuake users[5] would as far as to suggest there is no such thing as a "fair advantage" – trying to overcompensate for an inaccessible system by creating supposed advantages for disabled users may simply serve to confuse them.

[2] A discussion of alternative, more loosely coupled architectures is out of the scope of this paper but is currently being investigated.

[3] This does not preclude effects such as echos from being used to give the user an impression of their immediate environment.

[4] These differences, and their implications, will be discussed in a separate paper.

[5] There is a core set of roughly 15 blind and vision-impaired gamers that take part in our electronic discussion group.

3.2 Case Study: The EtherScan Radar

An example of the above phenomenon can be found in the behaviour of a navigation aid present in AudioQuake. The "ESR" warns players of nearby enemies and team mates using a RADAR-like metaphor: sounds emanate from the position of the enemy and have a gain and repetition speed proportional to the players' distance from them.

Originally it was conceived that allowing the ESR to "see" beyond walls and doors would help give blind players more chance of survival in the game. In reality, this feature – intended to help users – interfered with their sense of global navigation and caused them to track enemies behind walls, getting both stuck on obstacles and frustrated in the process.

Alternative indicators for enemies obscured by walls (such as muffled sound effects) could be used – and their effects may be investigated. However, the principle of conveying the required information in the simplest possible way and the avoidance of information overload [8] has been a prevalent (and successful) theme during the development and use of AudioQuake.

This result implies that mixing local and global navigation in an ad-hoc manner, within the same modality, can be confusing for users and detrimental to their ability in and enjoyment of games (and, thus, similar other activities).

3.3 Generalisation

A greater understanding of why certain pieces of information can be omitted from the auditory representation of the visual scene could be useful. This would enable us to construct a model of how people with certain disabilities (sight loss in this case) navigate, highlighting the types of information they require to do so effectively. From such a model, it would be possible to determine how this required information might be provided in this and other settings. Currently the implementation of navigation aids in AudioQuake serve as an empirical model. However, creating a more abstract and general version is an important goal and further work is being carried out in this area.

4 Serialisation and Prioritisation

A common generic approach to increasing the accessibility of a system is to (a) create a method for serialising its output and (b) put measures in place to ensure that the linearised output remains understandable (this may involve prioritising the rendering of parts of the output based on user needs). Examples of this are the way in which screenreaders interpret HTML pages and the work of the LAMBDA project [9]. This approach is popular and has a number of advantages, some of which are described below.

- It reduces multidimensional problems into single-dimensional problems, which may be easier to understand or at least display (the two most popular accessible output formats are speech and 1-dimensional Braille displays – both of which are linear).

- In some cases it is the more cost-effective and achievable approach, especially in areas such as web accessibility where retrofitting is more attractive to companies than redesigning their web presence[6].
- Sometimes it is the *only* known effective way of conveying information in an accessible form.

However, there are also some significant disadvantages to serialisation, which are highlighted below.

- Multiplexing nominally multidimensional data so as to render it via a 1-dimensional output medium may be cognitively demanding for users.
- Though it could be the most economically attractive approach, requiring less re-working of existing systems to implement, it may well not provide as sophisticated a level of accessibility as other methods[7].
- On its own it simply puts a (potentially large) amount of information into a format that can be "read". This does not necessarily make the information easier to understand (problems with navigation through the data may well occur). Other techniques are needed to complement serialisation to prevent information overload[8].

4.1 Domain-Specific Factors and Solutions

The above points mention some systems that the user can interact with largely at their own pace. However, many 3D computer games (and similar systems) pose a number of additional challenges:

- They are often fast-paced and time-critical.
- They generate a large volume of (mostly visual) information.
- This information spans multiple domains (spatial, strategy, storytelling and is presented in parallel.

AudioQuake does have some basic characteristics in common with hyper-stories such as AudioChille [11], however it is much more like other audio/accessible first-person-shooter games such as Terrafomers [2] and Demor [1], though unlike the latter it is targeted firmly at users who can only access commodity computing hardware.

[6] e.g. Screenreaders can linearise HTML tables with little effort required on behalf of the content producer. The planning agent approach [10] requires the author to undertake some extra work in order to work most autonomously (the adoption of newer mark-up standards would negate the need for a lot of this work, but that alone incurs significant effort).

[7] e.g. The planning agent approach [10] is capable of providing significantly more power and flexibility to the user than the serialisations of tables produced by a screenreader, but to work most autonomously it requires that the site be coded in XML (partly due to the semantic reinforcement afforded by meta-data).

[8] The LAMBDA project, for example, makes use of a hierarchical structure-exploration mode, for example.

One major difference, however, is that it is based on technology (Quake) not originally designed for non-sighted or otherwise disabled players[9]. Though this technology's architecture provides a solid base for building accessible games, the challenges listed above are potentially much more pronounced due to the fast-paced and unforgiving nature of the gameplay. Our approach to managing these issues is centred around the following key ideas:

– We imagine that there is a certain amount of *bandwidth* available for sending data from the computer system to the user (this may vary based on user capabilities).
– Information is *streamed* from the computer to the player. Some information is more urgent than other information and must therefore be rendered in an appropriate order.
– Different *domains* of information may be sent on different "channels"; these roughly correspond to different output devices/types of output – i.e. the system is multimodal.

We present a number of complementary techniques for effective serialisation.

– *Periodic Rendering by Priority* - The main tasks carried out by modern games and 3D applications run continuously. In one "tick" the screen is updated and physics rules, AI and gamecode are executed to update the user on the current state of the virtual world. Though it is appropriate to render visual, and some auditory, information at a high rate, this is not always necessary. For example, the player doesn't need to be constantly reminded of the locations of powerups (weapons, health and other such items). Indication of their position can be given periodically, as opposed to continuously.
– *Sub-Domain Prioritisation* - Information in a given domain may have varying importance. For example, enemies further from the player, or out of weapon range could be rendered with a lower priority than those which are within range. Similarly for powerups.
  Different schemes for prioritisation within each domain/stream were developed and found to be of use to AudioQuake players[10]
– *Multimodality across Domains* - The benefits of multimodal interfaces have been discussed extensively [12,13]. Rendering each stream to a different output device may not be possible, as there are a limited number of commodity output devices available. However, rendering all navigational information using non-speech audio and presenting communication between users using text-to-speech seems to have enabled AudioQuake players to understand the two sound-based streams separately, as intended.
  The use of haptic feedback (from Braille Displays to force-feedback input devices) could further separate out streams and enable users to interact more comfortably with the system, as demonstrated by other research [14].

[9] Fortunately, it was designed well, which enables us to implement accessibility – alternative rendering – at the correct level.

[10] This should not be confused with the filtering techniques described above – even after filtering out surplus information, some prioritisation may be necessary.

– *Cross-Domain Prioritisation* - When the pace of events in the game increases, the player could require more information to complete certain tasks. For example, when interacting with a number of enemies, they should ideally be informed of the position of all enemies (as opposed to just the closest) and perhaps any nearby powerups which could assist them.
  This has not yet been implemented but feedback from users indicates it could be worthwhile. The general idea of load-balancing in computer networks or the grid [15] could well be applied in this situation – with the goal of sharing output bandwidth amongst modalities and devices in such a way that the user does not become overloaded with information.

More implicit, techniques for improving bandwidth usage and immersing the player more in the game are out of the scope of this paper and will be discussed separately.

5 Links to Education

There is a growing interest in the potential usefulness of game-like technology for education [16,17]. This work aims to promote inclusion in a number of ways.

– *Application to Other 3D Engines* - The techniques developed are quite generic and could be employed elsewhere (for games or other applications).
– *Collaboration* - Though out of the scope of this paper, AudioQuake promotes accessible online collaboration, promoting integration with sighted students who may be using the mainstream version of the same technology.
– *Uses in Other forms of Navigation* - Mobility training and the navigation of complex data structures are similar tasks in many respects and could make direct use of the navigation aids described here.
– *Direct Educational Uses* - As well as the potential applications listed above, some more direct benefits of accessible 3D engines and games exist. They can be used as practical material in the teaching of programming, networking and even the development of important algorithms related to AI/machine-learning and searching [18].

6 Conclusions

We have discussed a number of issues that are key to providing accessibility to 3D environments such as computer games. Many of these challenges – relating structure/space to the user; presenting information that the user needs, quickly – are common to a number of accessibility and usability problems. We hope to generalise the currently domain-specific solutions to other areas.

References

1. Velleman, E., van Tol, R., Huiberts, S., Verwey, H.: 3d shooting games, multimodal games, sound games and more working examples of the future of games for the blind. Lecture Notes in Computer Science **3118** (2004) 257–263

2. Westin, T.: Game accessibility case study: Terraformers - a real-time 3d graphic game. In: The Fifth International Conference on Disability, Virtual Reality and Associated Technologies. (2004)
3. Archambault, D., Burger, D.: From multimodality to multimodalities: the need for independent models. In Constantine Stephanidis, L.E.A., ed.: Proceedings of the UAHCI'01 conference "Universal Access in HCI - Towards an Informatino Society for All", New-Orleans, United States of America (2001) 227–231
4. Hanson, V.L., Richards, J.T.: A web accessibility service: update and findings. In: Assets '04: Proceedings of the 6th international ACM SIGACCESS conference on Computers and accessibility, New York, NY, USA, ACM Press (2004) 169–176
5. Pontelli, E., Xiong, W., Gupta, G., Karshmer, A.I.: A domain specific language framework for non-visual browsing of complex html structures. In: Assets '00: Proceedings of the fourth international ACM conference on Assistive technologies, New York, NY, USA, ACM Press (2000) 180–187
6. Brewster, S., Brown, L.M.: Tactons: structured tactile messages for non-visual information display. In: CRPIT '04: Proceedings of the fifth conference on Australasian user interface, Darlinghurst, Australia, Australia, Australian Computer Society, Inc. (2004) 15–23
7. Smith, A.C., Cook, J.S., Francioni, J.M., Hossain, A., Anwar, M., Rahman, M.F.: Nonvisual tool for navigating hierarchical structures. In: Assets '04: Proceedings of the 6th international ACM SIGACCESS conference on Computers and accessibility, New York, NY, USA, ACM Press (2004) 133–139
8. Brewster, S.A.: Using non-speech sound to overcome information overload. Displays **17** (1997)
9. LAMBDA Project: Linear Access to Mathematic for Braille Device and Audiosynthesis. http://www.lambdaproject.org/ (2005)
10. Pontelli, E., Son, T.C.: Planning, reasoning, and agents for non-visual navigation of tables and frames. In: Assets '02: Proceedings of the fifth international ACM conference on Assistive technologies, New York, NY, USA, ACM Press (2002) 73–80
11. Sánchez, J., Sáenz, M.: 3d sound interactive environments for problem solving. In: Assets '05: Proceedings of the 7th international ACM SIGACCESS conference on Computers and accessibility, New York, NY, USA, ACM Press (2005) 173–179
12. Latoschik, M.E.: A user interface framework for multimodal vr interactions. In: ICMI '05: Proceedings of the 7th international conference on Multimodal interfaces, New York, NY, USA, ACM Press (2005) 76–83
13. Baljko, M.: The information-theoretic analysis of unimodal interfaces and their multimodal counterparts. In: Assets '05: Proceedings of the 7th international ACM SIGACCESS conference on Computers and accessibility, New York, NY, USA, ACM Press (2005) 28–35
14. Sjöström, C.: Using haptics in computer interfaces for blind people. In: CHI '01: CHI '01 extended abstracts on Human factors in computing systems, New York, NY, USA, ACM Press (2001) 245–246
15. Berry, R.: Performability and grid in an on demand world. In: Grid Performability Modelling and Measurement, Edinburgh, National eScience Centre (2004)
16. NESTA, EA: Futurelab. http://www.nestafuturelab.org/research/teachingwithgames.htm (2005)
17. Gee, J.P.: What video games have to teach us about learning and literacy. Comput. Entertain. **1**(1) (2003) 20–20
18. Laird, J.E.: Using a computer game to develop advanced ai. Computer **34** (2001) 70–75

Access Invaders: Developing a Universally Accessible Action Game

Dimitris Grammenos[1], Anthony Savidis[1],
Yannis Georgalis[1], and Constantine Stephanidis[1,2]

[1] Foundation for Research and Technology - Hellas (FORTH),
Institute of Computer Science (ICS), Human-Computer Interaction Laboratory,
Heraklion, Crete, GR - 70013 Greece
{gramenos, as, jgeorgal, cs}@ics.forth.gr
[2] Department of Computer Science
University of Crete

Abstract. This paper depicts the notion of Universally Accessible Games and presents the development of a related action game entitled Access Invaders. The design of the game's user interface which accommodates concurrently the needs of people with diverse abilities is described, along with the approach followed to adapt the game logic and content to achieve accessibility. In this context, the concept of Parallel Game Universes is introduced and suggested as a solution for the creation of multiplayer universally accessible action games.

1 Introduction and Problem Description

In the past few years, the accessibility of electronic applications and services by disabled people has become a topic of paramount importance at an international level. Based on the fundamental right of all people for access to information and services, and equal opportunities for employment and independent living, several governments and international political bodies have adopted legislative and policy measures for application software and Web accessibility (e.g., [1], [2]). One of the basic needs of most people, beyond working and independent living, is entertainment. Presently, computer games constitute indisputably one of the major related sources.

Unfortunately, computer games are usually quite demanding in terms of motor, sensor and mental skills needed for interaction control, while they often require mastering inflexible and complex input devices and techniques. These facts often render games inaccessible to a large percentage of people with physical (or situational) disabilities. So far, little attention has been paid to the development of computer games that can be played by all players, independently of their personal characteristics, requirements, or (dis)abilities. Furthermore, concerning human-computer interaction issues, computer games have fundamental differences from all the other types of software applications, for which accessibility guidelines and solutions are already becoming widely available.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, , pp. 388 – 395, 2006.

2 State of the Art in R&D and Application

In contrast to Web accessibility, up to now, relatively few efforts have been devoted to game accessibility. Currently, there are no related official guidelines or standards, nor any world-wide initiatives promoting game accessibility, and no related governmental or legislative actions. At present, the only support and design knowledge available to developers for creating accessible games is limited to some indicative approaches outlined in a White Paper of the Game Accessibility SIG of IGDA [3], as well as general-purpose guidelines for developing accessible software (e.g., [4,5,1]). From a technical point of view, two main general approaches have been adopted to address the issue of computer games accessibility:

1. Mainstream games are developed to be compatible with the use of assistive technologies, such as screen readers, mouse emulators or virtual keyboards.
2. Special-purpose games are created, optimally designed for people with disabilities, like audio-based games for the blind and switch-based games for the motor-impaired.

The first approach typically achieves very limited accessibility and suffers from low *quality in use*. The second approach, though being the most promising from a quality point of view, has two key drawbacks: (a) the cost of developing high quality games is prohibitive when the potential target group is limited; and (b) there is an evident hazard of segregation between able-bodied and disabled gamers, leading to potential social exclusion.

Currently, there are no computer games that can be concurrently played by people with different disabilities sharing the same computer. Recently, a few games for both visually-impaired and fully sighted players have been developed (e.g., *All inPlay*[1] card games*;* 3D shooter Terraformers[2]). Several audio-only games adopting the Space Invaders theme exist for the blind (e.g., a.Shooter[3], Sonic Invaders[4], Alien Outback[5]). All are single-player sound-only games, with no configuration capabilities. A game with both audio (in Japanese) and visual output is Space Invaders for the Blind[6] by Taito. A 3D Space Invaders game has been developed [6] which combines audio and visual interfaces with force. Regarding people with motor impairments and low vision, only *Alien Invasion*[7] is available. The game can be played with a range of standard and adaptive technology controls, and allows players to adjust the parameters of the major game elements, also offering large character sizes for low vision players.

---

1 `http://allinplay.com/`
2 `http://www.terraformers.nu`
3 `http://www.agame.org/en/aShooter/index.html`
4 `http://www.mytaras.com/sonicinvaders.html`
5 `http://www.espsoftworks.com/`
6 `http://homepage2.nifty.com/JHS/spi.html`
7 `http://www.arcess.com/aliens.htm`

3 Research and Methodological Approach

In order to overcome the limitations of previous approaches to game accessibility, the concept of of *Universally Accessible Games*[8] (UA-Games) has been proposed, primarily emphasizing game accessibility, but also putting forward the objective of creating games that are concurrently accessible to people with diverse abilities. UA-Games are interactive computer games that:

1. follow the principles of *User Interfaces for All* [7], being proactively designed to optimally fit and adapt to different individual gamer characteristics without the need of further adjustments or developments;
2. can be concurrently played among people with different abilities, ideally also when sharing the same computer;
3. can be played on alternative technological platforms and contexts of use using a large variety of devices, including assistive technology add-ons.

The potential impact of UA-Games in the upcoming Information Society is threefold: (a) they open up and enhance an entertaining social experience that would otherwise be unavailable to a significant percentage of people; (b) they allow for social interaction among people who may never have (or could have) interacted with each other; and (c) they considerably expand the size and composition of the potential market of the computer games industry.

3.1 From Space Invaders to Access Invaders

Action games constitute a real challenge for Universal Access. They have highly dynamic content, since they comprise many different moving objects with alternative characteristics, they usually require complex controls and are based on reflex-based reacting, a fact that can render them particularly hard for those who have difficulty or cannot use their hands or eyes. Also, devising ways for two players with different disabilities to play, cooperatively or against each other, the same action game, is a very challenging research issue.

In this context, a universally accessible version of the classic action game *Space Invaders* by TAITO has been designed and developed. The new game, named *Access Invaders,* achieves Universal Access by supporting alternative input / output modalities and interaction techniques that can co-exist and cooperate in its user interface, combined with tailorable player profiles and game content. The game is highly customizable and supports the creation and use of unlimited user profiles. Each game parameter can be adapted both based on the player's profile and the current game level. Non-visual gameplay is also supported. In this case, full acoustic rendering of game information is provided through spatial audio and a built-in screen reader. Multi-player games are available, where people with different (dis)abilities can play cooperatively, sharing the same computer. In this case, the game's interaction parameters can be independently adjusted for each player. An unlimited number of concurrent players is supported. Access Invaders is available for download as freeware at `http://www.ics.forth.gr/hci/ua-games/access-invaders`

[8] `http://www.ics.forth.gr/hci/ua-games`

3.2 Game Design

A prerequisite for creating UA-Games is *inclusive design*, which implies that the requirements of the broadest possible population should be taken into account during the design phase. For designing Access Invaders, the following user categories and respective requirements were considered: (a) people with hand-motor impairments; (b) blind people; (c) people with deteriorated vision; (d) people with mild memory / cognitive impairments and novice players; (e) people belonging in more than one of the previous groups. To elicit user requirements, three basic sources were used: (a) relevant bibliography; (b) interviews with experts and representatives of target user groups; and (c) observation of potential users while playing related computer games.

In order to be able to effectively accommodate the particularly broad spectrum of diverse interaction requirements, the game was firstly designed in a context-independent way, i.e., at an abstract level, minimizing references to specific interaction modalities, metaphors, techniques or devices. Additionally, it was necessary to appropriately map the abstract design elements to low-level, physical, interaction styles, meeting the requirements of each target user group. In this context, accessibility barriers related to the game's interface, but also barriers stemming from the game's content and rules were identified, as along with possible design strategies for overcoming them. For these purposes, the Unified Design method [8] was adopted, reflecting a procedural design discipline of abstract task definition with incremental polymorphic physical specialization.

Based on the high-level design, indicative electronic prototypes of the game were developed, showcasing alternative interactive properties of its user interface for the different target user groups. These prototypes were informally evaluated with the stakeholders that participated in the design process using the *thinking aloud method* [9]. During the evaluation, participants were prompted to express their thoughts, comments and feelings. The outcomes of this process aided in validating, correcting and updating design decisions, as well as in developing new ideas for improving the accessibility of the final game.

3.3 Adapting the Game to the Player's Needs

The main mechanism of game adaptation are *profiles*. Once a profile is selected the game's interaction and content adapt to its characteristics. Thus, in practice, the game is viewed by players as a collection of different games integrated into one, from which they can select the one that is most suitable for them. On the other hand, from a developer's point of view, it is a single application with unlimited alternative manifestations. *Access Invaders* currently offers seven predefined profiles but new ones can also be created. Some indicative examples of the game's adaptation capabilities are provided in Fig. 1.

Adapting the User Interface. The game's interface can be controlled using any, or all, of the following devices: keyboard, mouse, joystick, game pad and binary switches. The interface includes two basic interaction objects, i.e., menu and text entry. Both can be manipulated with any of the aforementioned devices, while the font size and family is configurable. Furthermore, automatic scanning is provided for

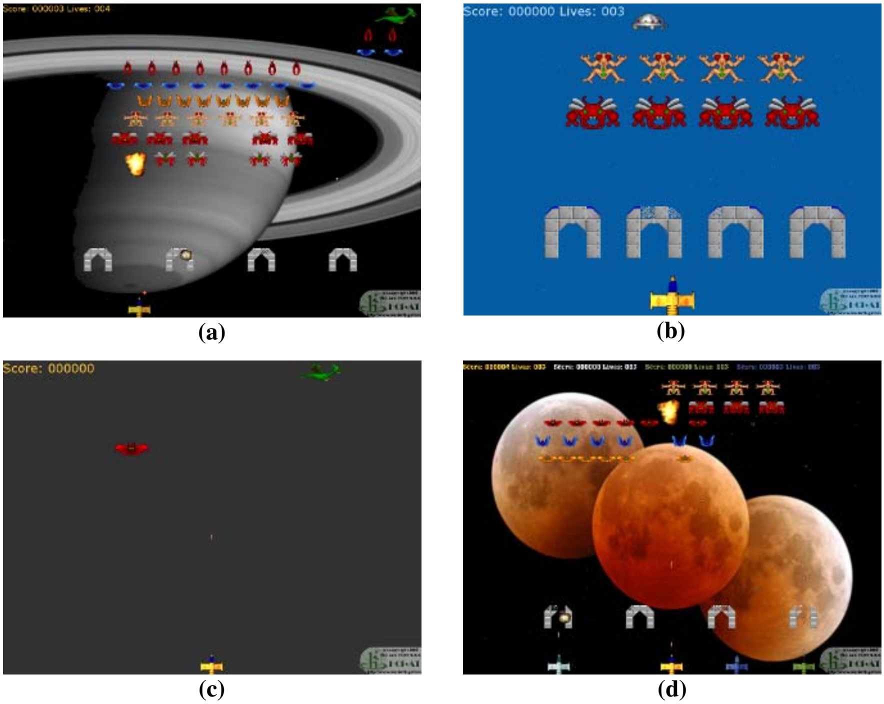


Fig. 1. Examples of the user interface and content adaptation capabilities of *Access Invaders*: (a) typical single-user game, (b) content enlargement and visual simplification for players with low-vision, (c) audio-based game for blind players, (d) cooperative game of multiple players with different (dis)abilities

people who can use only a single switch. For supporting non-visual interaction, the objects are augmented with speech output capabilities. The player's spaceship is controlled through three commands("move left", "move right" and "fire"), which can be issued using any device. In order to achieve accessibility for people with limited motor functions, the game can also be played using only two, or even a single command. Furthermore, two more features are supported (combined into one): (a) pausing the game, and (b) accessing an "in-game" menu with functions for resuming, quitting or changing parameters of the current session. Typically, these features are accessed through additional dedicated controls (e.g., Escape key, right mouse button). Nevertheless, in the case of limited motor control this is not possible. In this case, the menu is activated if all the available commands are concurrently issued for a certain amount of time (e.g., if the player keeps the "change direction" switch pressed for more than 3 seconds).

Adapting the game content and logic. As a result of an analysis of the currently available accessibility guidelines for applications and games, and after several hours of play testing with end-users, a number of adaptation categories were identified which are outlined in Table 1. Alternative combinations of these adaptations can serve the accessibility needs of each of the addressed target user groups.

Table 1. Overview of game content and logic adaptation categories

Category	Example
	Game Content
Speed	Make individual game elements, or the gameplay, faster or slower.
Quantity	Change the number of aliens and shields that are present.
Size	Make the game elements bigger / smaller.
Layout	Alter the absolute / relative position of game elements.
Firepower	Supply the aliens / player with faster / slower and more/less powerful weapons or stop them from firing.
Visual complexity	Use of backgrounds and graphics with large solid colour areas.
Contrast	Use of colour combinations that provide high contrast.
Sound	Association of spatial feedback to game elements through 3D sound.
	Game Rules
Interaction among game elements	Set which type of aliens can destroy the player's spaceship and vice versa, and whether the spaceship's bullets can pass through the shields or collide with them.
Analogue vs. digital control	Select whether the aliens / spaceship move in discrete or continuous positions.
Hints	Make the player's task easier by providing added-value information, e.g., audio radar, visualisation of the path of objects, oral descriptions.
Stamina	Provide the player with additional "lives", make the spaceship more resistant to incoming fire and the aliens less.

3.4 The Concept of Parallel Game Universes

Multiplayer action games constitute an open research challenge, since it is not only the interface that changes, but also the game's content and rules. This practically means that two (or more) people should be able to play the same game, being aware of each other, while at the same time each one follows different rules and perceives distinct content. A possible solution to this problem is to allow each player to play in a different "game universe" and then somehow project each universe to the other(s). The term "game universe" is used to denote an instance of the game after it has been adapted to suit the requirements and needs of a particular player. For example, the alternative profiles of Access Invaders could be considered as different game universes.

As an illustrative example, consider the following situation. Two friends want to play the game together. One of them (Player X), due to severe motor-impairments, can use only a single switch. A manageable difficulty level includes a small group of aliens that move slowly and fire very scarcely, while the player's bullets do not collide with shields. The second Player (Y) does not have any impairment and in order to for the game to be challenging enough, he wants to confront numerous fast, fire blazing aliens. So, in case the two players attempt to share the very same game, if this is adapted to the first player, then it will be rather boring for the second, while if it is adapted to the second player it will be extremely difficult - if not impossible - for the first. A novel approach, following the idea of *Parallel Game Universes*, is to merge the two distinct game versions into one. Thus, in this new version of the game, two groups of aliens will exist: a big, fast and powerful which can destroy and be destroyed only by Player Y, and a small, slow and quite harmless that plays only against Player X. The bullets of each player will not affect the aliens fighting against the other, while Player Y's bullets will collide with the shields, and Player X's will not.

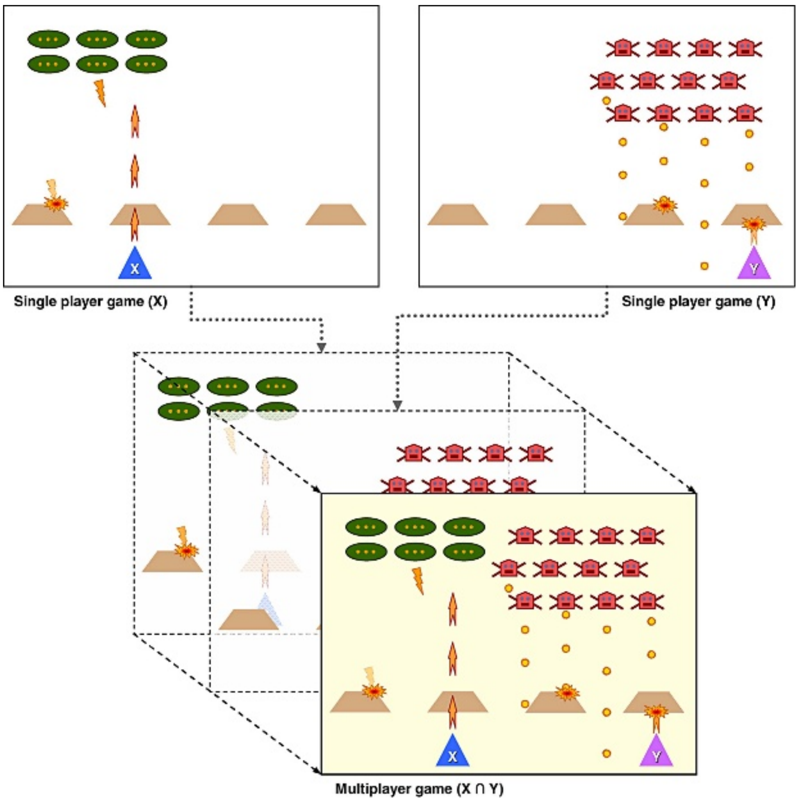


Fig. 2. Example of universally accessible multiplayer gaming through Parallel Game Universes

4 Future Work

In the short term, future work includes the support of tactile output trough a Braille display and a force feedback joystick and stylus. Furthermore, an interactive application for editing user profiles and related game levels is under development. In the medium term, planned research and development work can be divided in two different, but highly interrelated, areas: (a) *interactive profile selection*, through which the game will help players in selecting a profile and game parameters that best suit their needs; and (b) *dynamic gameplay adaptation*, which includes monitoring the player's actions and dynamically adjusting the gameplay to better match the player's skills. Longer term work entails further research and elaboration of the concept of *Parallel Game Universes*, since the current version of Access Invaders supports only parallel universes that can be instantiated in the same computer. In this context, the main objective is, on the one hand, to further experiment and study the potential and the limits of this approach, while, on the other hand, to implement and evaluate prototypes of distributed game universes.

References

1. Access Board (2000). Electronic and Information Technology Accessibility Standards. Retrieved September 14, 2004, from: http://www.access-board.gov/sec508/508standards.htm
2. European Commission (2002). eEurope 2005: An information society for all. Retrieved September 4, 2004, from:http://europa.eu.int/information_society/eeurope/2005/index_en.htm
3. GA-SIG (2004). White Paper: Accessibility in Games: Motivations and Approaches. Retrieved September 4, 2004, from:http://www.igda.org/accessibility/IGDA_Accessibility_WhitePaper.pdf
4. W3C (2004). W3C-WAI resources: Guidelines. Retrieved September 14, 2004, from: http://www.w3.org/WAI/Resources/#gl
5. ISO (2003). ISO/TS 16071:2003: Ergonomics of human-system interaction -- Guidance on accessibility for human-computer interfaces. International Standards Organisation, Geneva, Switzerland.
6. Mccrindle, R. J. and Symons, D. (2000). Audio Space Invaders. In Proceedings of the 3rd International Conference of Disability, Virtual Reality & Associated Technologies, Alghero, Italy, 2000, 59–65.
7. Stephanidis, C. (2001). User Interfaces for All: New perspectives into Human-Computer Interaction. In C. Stephanidis (Ed.), User Interfaces for All - Concepts, Methods, and Tools (pp. 3-17). Mahwah, NJ: Lawrence Erlbaum Associates.
8. Savidis, A., Stephanidis, C. (2004). Unified User Interface Design: Designing Universally Accessible Interactions. International Journal of Interacting with Computers, Elsevier, Issue 16, 243-270.
9. Nielsen, J. (1993). Usability Engineering. Boston, MA: Academic Press, pp.195-199.

Internet and Accessible Entertainment

Morten Tollefsen and Are Flyen

Media Lunde Tollefsen AS, Tevlingveien 15, 1081 Oslo, Norway
{morten, are}@medialt.no

Abstract. Young people with disabilities should be entitled to use technology for entertainment, but research and development in IT and the disabled has not typically focused on pleasurable activities. It is therefore important to set up guidelines that facilitate accessibility, develop good examples of accessible computer games, and not least, utilize the great potential of technology to include persons with different requirements.

1 Background

The UPS-project [1] started in October 2003 and runs for a period of three years. The project is aimed at adapting and developing entertaining software for young people with disabilities. The target group includes those with mental disabilities, or a combination of severe motor handicaps and perceptual losses. The work can be divided into three main areas:

- Adaptation of standard software
- Development of software designed to meet the needs of the target group
- Design of a website with information, game tests and free, digital materials

The project has focused on standard software, and how this can be inclusive, allowing disabled young people to use computer games popular among friends and family. However, we found during the course of our work, that it would be necessary to develop software for the most disabled users. This software should contain the best aspects of standard computer games and entertaining products. With this perspective in mind, software development for the target group, exemplified by an internet based game, is described below.

2 State of the Art

2.1 Entertainment and Simple User Interfaces

Research and development in the field of IT and the disabled has focused on education rather than leisure. Computer games, media players and other entertaining software are important for many people, and these products are typically developed without an educational motive. Entertainment is important, and of course just as important for disabled persons as for others (in some cases perhaps even more

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 396 – 402, 2006.

important). Educational software is of course useful and in some cases even fun! The opposite is equally true: entertaining software may provide good motivation e.g. for learning assistive devices and standard computer skills. An example of this is found in [3], and Mr. Phillips writes: “Whenever I look at a computer, I think of all the fun to be had, not how I'll be able to write essays or operate the lights in my room.” And “Yes, games like Warcraft III and Age of Mythology can be fast paced, but it was my desire to play such games that made me the highly proficient scanning user I am today. Gaming taught me speed”.

Several standard games can be used by the disabled, but there is also a need for simple games developed for fun only, and for the most disabled users. “Easy Games” (`http://www.leripa.se/easygames.asp`) and games found at `http://www.arcess.com/` are examples of entertaining software for our target group. Other products, more or less educational, do exist but our experiences indicate that these products are typically either too childish or difficult to operate for young persons with the largest disabilities. Another important fact is that games are not typically localized into small languages like Norwegian.

There are few accessible net based games. Some research and development has been done for the visually impaired [4]. As described below, a net based game was desirable in the UPS-project. We have not found such games with built-in scanning etc.

2.2 Standards and Guidelines

The development of guidelines for developing accessible games [2] has been an important activity in the UPS-project, and this work was organized as an international co-operation in 2005. Guidelines are available at `http://gameaccess.medialt.no/guide.php`. We have implemented the most important guidelines for our target group in the product, and even tried to allow for users with other needs, e.g. that blind persons are able to use the game.

The game is developed using Macromedia Flash, and Macromedia Flash guidelines “Best Practices for Accessible Flash Design” [5] have been helpful. Supplementary product information like frequently asked ques-tions is found on standard web pages, and WCAG 1.0 (`http://www. w3.org/wai/`) are emphasized.

3 Research and Methodological Approach

3.1 What Do the Users Want?

We did not want to spend a great deal of effort developing a product which was not appropriate for the target group. This is quite possible when developing a product for users with limited communication possibilities. Our approach to discovering what the users actually wanted was to arrange two brainstorming sessions/seminars with several experts. We also mapped the preferences of pupils at a special school (Haukåsen School). Although interests vary greatly (some prefer guns, others dolls) the conclusion from this work was: “internet and music”.

3.2 Best Guess Design

A participatory design approach has been used when developing the product. However experts (teachers, pedagogues etc.) have in most cases represented the young disabled players. When designing for persons with mental disabilities, or a combination of severe motor handicaps and perceptual losses, it is necessary to build on earlier experiences, both when designing tasks and user interface. Best guess design is therefore an adequate description of the design process used when developing HeiPipLerke.

Previous work and experiences (e.g. from developing electronic picture books) suggested four important keywords for developing a new entertaining product:

1. Entertaining
2. Simple (not childish)
3. Accessible
4. Flexible

3.3 User Testing

Best guess design must be succeeded by user testing. In our case user tests should answer two important questions:

1. Is HeiPipLerke entertaining?
2. Is the product accessible and easy to use?

User testing has been performed at three Norwegian centres of competence and a school for the disabled. An interview scheme was used; however most of the findings are based on observations from teachers/helpers. In addition to direct user testing it has also been possible to send feedback from the web.

3.4 Development Tools and Accessibility Support

Although the conclusion from brainstorming sessions, expert views and the mapping of user preferences indicated an internet product, we still had to consider distribution media and which operating systems to support. It is probably easier to allow for accessibility when using standard programming languages or tool kits for disabled than to use the most popular web development tools. New internet technologies do also focus on accessibility and universal web accessibility seems to be constantly more important. We believe that in the future even more accessibility features will become available, e.g. in Macromedia Flash.

Using internet has several advantages: support for different platforms, no client installation, only one source to maintain, and less time spent on support. Three different technologies were considered: Macromedia Flash, Java (Applets) and traditional HTML/Javascript with server-side interaction. The final choice was Macromedia Flash (Macromedia is now known as Adobe). The reasons for choosing Macromedia Flash were:

- Flash is suitable for built-in tools for creating graphics and animations
- Using mp3-files for music, speech and sound-effects is straightforward
- Accessibility support is fairly good

- Easy distribution to different medias like internet and CD-ROM (offline edition)
- Visual appeal and short development time

4 HeiPipLerke (Meadow Pipit)

The game can in short be described as a music composer adjusted to meet the requirements of the target group. When the game is started a short introductory screen is displayed while the game loads and brings up the main menu. The main menu contains three choices: start the game, setup and introduction.

When starting the game for the first time both automatic scanning and speech (audible scanning) are enabled by default. In the setup screen it is possible to enable or disable scanning and speech, set scanning interval, change the colour-scheme for the entire game, select touch screen support etc. Choices are saved and automatically used when the game is started again.

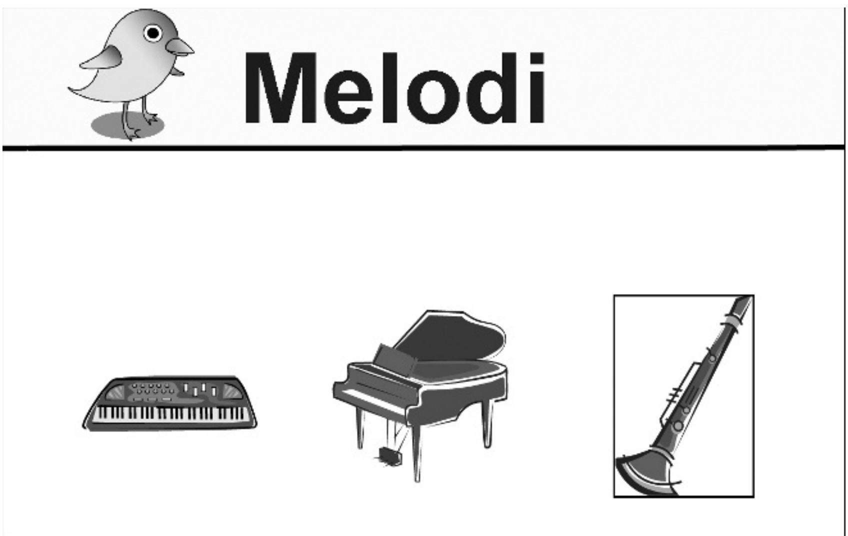


Fig. 1. This screenshot shows one of three possible instruments for the melody with focus

A set of boxes is shown when the game is started. Each box can be described as a music collection. When a music box is selected, the next three screens display choices for three different categories: melody, beat and bass. In each category there are three different instruments to choose between which gives 3 x 3 x 3 = 27 different songs that can be created from each initial box. More music boxes can easily be added at a later time to extend the game and give more variations.

When the "composition" is complete a simple animated show is displayed while the song is played. Each song is 40 seconds long. After playing the song there are three choices: repeat playback, return to the main menu, or send the song to a friend. The latter is a form where the players name and a friend's e-mail can be filled in. An e-mail is then posted to the address with a link to a website where the song can be played using an accessible media player.

5 Findings, Challenges and Solutions

Macromedia Flash was the most suitable development tool for our purpose but we faced some challenges with respect to accessibility. Some problems are described in [6]. Macromedia has however added accessibility features in MX and version 8. The main features are naming and description of all objects and tab-indexing. These features are not appropriate when developing an accessible game, and our most important findings are discussed below.

5.1 Scanning and Built-In Speech/Sound

In HeiPipLerke scanning and speech are required functionality. This functionality is not part of Flash, and therefore had to be implemented. The result of this work is techniques to add visual- and audible scanning and speech with or without scanning enabled. When using built-in speech, scanning functionality has to allow for the application messages before focusing on specific selectable objects. This is the case e.g. when asking the user to select a melody. The first melody cannot be selected/ played before the guidance is spoken.

5.2 Mouse over and Touch Screens

One important feedback from the expert testers was the “mouse over” function (or object focus) e.g. an instrument starts to play when pointing on it with the mouse. This functionality is important when learning assistive devices, and is not typical in standard games.

A somewhat different challenge was reported for persons using a touch screen. When pointing to an object the object was focused. But when pressing a finger on the object, it was immediately selected i.e. the instrument did not play an example before being selected. The solution we chose was to add an option in the touch screen setup. When pressing the object for the first time it will be focused i.e. start to play and when pressing a second time the object will be selected.

5.3 Screen Readers, Color Schemes and Focus Rectangle

Although HeiPipLerke is not primarily designed for blind and visually impaired users, our goal was to make the product as accessible as possible. Flash has built in options for describing elements (name, description) and defining tab-index. These accessibility options make it possible to use the product, although the user may have to adjust the standard settings in e.g. a screen reader. One example is that it is an advantage to disable the virtual PC cursor in Jaws.

Text can be read as text (not rendered as graphics). However screen readers like Jaws will typically focus on the text elements when pressing Tab. This is not desirable in a game, and user testing indicates that a better solution is graphics with appropriate tagging.

At least two different high/low-contrast colour-schemes are needed to support visually impaired users (black on white, and white on black). Most of the objects could easily be given a new colour in runtime but a few objects involved much more work. The standard focus rectangle in Flash was reported to be a major problem. By default this rectangle is too indistinct (thin, yellow frame). On some objects e.g. form fields the default color of the rectangle is green. The color as well as the size should be user definable. A distinct focus rectangle is important for persons with a visual impairment. It is also a requirement for other disabled users. To solve this problem the standard style sheets needed to be redefined. This resulted in a lot of work and research because the colour scheme/stylesheet was not appropriately documented by Macromedia.

5.4 Browser Compatibility

HeiPipLerke has been tested with the most recent versions of Internet Explorer, Opera and Mozilla Firefox. Getting focus on the Flash application is a problem in all browsers. The Flash application has to be focused before the user is able to use the keyboard. Microsoft Internet Explorer proved to be the best browser (even if an extra Tab has to be pressed).

5.5 Forms and Built-In Scanning

From the very beginning user scanning was implemented in all screens. These screens include two forms: Setup and Send mail. User testing resulted in disabling scanning in these forms. An on screen keyboard can be used to enter text (e.g. mail address); however many in the target group are not typically able to do this without a helper. An option could have been to enable scanning with an interruption when text is entered.

5.6 List of Mail Recipients

Send the music to a friend is a popular function. Originally we did not plan a "My friends mailing list". This is a very reasonable thing to add, and will make it much easier to send music for the target group.

6 Future Plans

HeiPipLerke is designed to work as an example of an accessible net based game and as a free entertaining product for young disabled persons. Feedback from experts and users indicates a need for products with similar user interfaces, both in applications as simple as HeiPipLerke and in more complex applications.

Localization: Music is an international language. It is therefore easy to translate HeiPipLerke. In addition to translating some text pages it is necessary to record corresponding sound files.

More contents: Music can easily be added and improved.

Hearing impaired: Because some persons in the target group of this product use sign language it is appropriate to include sign language in addition to speech. This can be done using a free Norwegian dictionary. Free sign language videos are available in other languages as well, and even this functionality should be straight forward to translate.

New games: The techniques and programming tools developed for this game can be used in other internet applications. The need for such applications is evident. The challenge is to raise funds, establish free programming resources (e.g. students) or develop commercial business.

References

1. Tollefsen, M., Scheidegger, D.: Computer Games For Young Persons With Learning Disabilities, CSUN 2005
2. Tollefsen, M., Lunde, M.: Entertaining Software for Young Persons with Disabilities, ICCHP 2004
3. Phillips, M.: Function follows fun: How games helped me to develop my switch skills, CSUN 2005
4. Velleman, E., van Tol, R., Huiberts, S., Verwey, H.: 3D Shooting Games, Multimodal Games, Sound Games and More Working Examples of the Future of Games for the Blind, ICCHP 2004
5. Regan, B.: Best Practices for Accessible Flash Design, Macromedia 2005
6. Jordan, M.: Strategies for building accessible learning web applications and games with Macromedia Flash MX, CSUN 2005

Guidelines for the Development of Accessible Computer Games

Roland Ossmann and Klaus Miesenberger

University of Linz, Institute Integriert Studieren
Altenbergerstr. 69, 4040 Linz, Austria
roland.ossmann@students.jku.at, klaus.miesenberger@jku.at

Abstract. Games are very important for learning, teaching, entertainment, inclusion. But they are of the most challenging applications concerning accessibility, and usability for people with disabilities. Especially in the context of playing together or in groups equal access is critical. In this paper we will present first attempts to define games accessibility guidelines helping game developers to design their products in a way that assistive technologies can interact with the game interface and that the parameters of usage can be adapted to the needs of people with disabilities.

1 Introduction

Computer games have become a major part in child and youth culture, but they are also played by adults. To give people with disabilities the chance to have access to these games, new methods and tools have to be developed. One possibility is to give game developers guidelines, with rules and hints how to develop accessible games.

This paper will present the work done in a working group which aims at starting an international co-operation to develop guidelines for games accessibility. The paper (in its final form) will a) the state of the art in research towards guidelines in games accessibility, b) present the state of the guidelines developed in the group c) present the system used to work on these guidelines at international level d) present other activities which should support developers in making their games more accessible ("Active Games Accessibility").

2 State of the Art

There are two main research projects in the area of guidelines for games accessibility. Project one is from IGDA, the International Games Developer Association. Inside IGDA is the Games Accessibility Special Interest Group (GA-SIG), who has published a whitepaper [1] about games accessibility, and one chapter is about rules and hints for game developers.

The other main project is done by the Norwegian IT company MediaLT. MediaLT has developed a set of guidelines [2], which were the basis for further development of out guidelines. Furthermore MediaLT is partner in our project.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 403 – 406, 2006.

3 Research

With the guidelines from MediaLT, the rules and hints form GA-SIG and out own ideas, we developed our own guidelines and published them as a web page [3]. A web page was decided to make the GL accessible to everyone who want to bring in new ideas or want to help making the existing GL better.

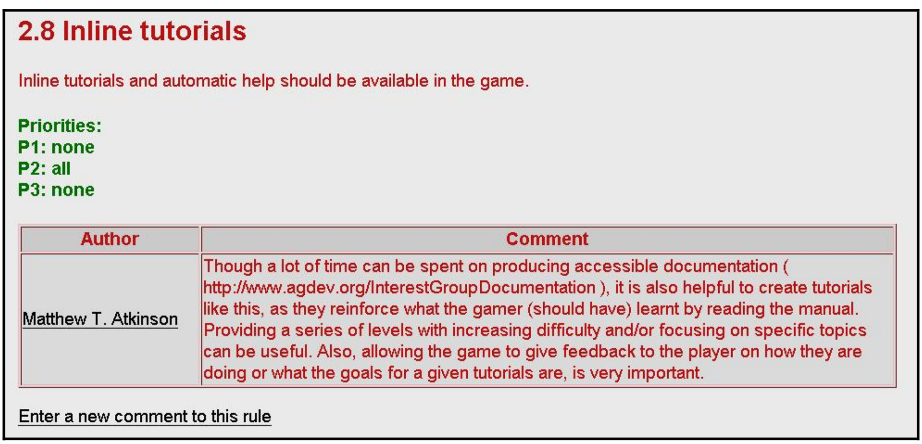
These guidelines have five main categories:

- level/progression
- input
- graphics
- sound
- installation and settings

The guidelines have, beside the rules itself, a categorisation in three classes of priorities:

- Priority 1 – Must have
  Must have means, that it is absolutely necessary for the listed group of gamers. Otherwise the game is not accessible for them.
- Priority 2 – Should have
  Should have means, that it is a big help for the listed group of gamers. The game is accessible without that point, but with that point, the game is easier to learn or the fun factor is higher.
- P3 – May have
  May have means, that it is a help of feature for the listed group of gamers. The game is accessible without that point.

Furthermore there are four groups of disabilities: visual, auditory, mobility and cognitive disabilities. These disabilities are allocated to the priorities, e.g. one rule can have priority 1 for visually impaired people and priority 3 for auditory impaired people.


2.8 Inline tutorials

Inline tutorials and automatic help should be available in the game.

Priorities:
P1: none
P2: all
P3: none

Author	Comment
Matthew T. Atkinson	Though a lot of time can be spent on producing accessible documentation (http://www.agdev.org/InterestGroupDocumentation), it is also helpful to create tutorials like this, as they reinforce what the gamer (should have) learnt by reading the manual. Providing a series of levels with increasing difficulty and/or focusing on specific topics can be useful. Also, allowing the game to give feedback to the player on how they are doing or what the goals for a given tutorials are, is very important.

Enter a new comment to this rule

Fig. 1. Screenshot of the Guideline Tool

In November 2005, a workshop about games accessibility was held, and in this workshop the further approach was decided. The future development of the GL is split up in two groups, a main working group and a board group for decisions and quality control. The communication works over mailing lists and the possibility to post comments to each rule of the guideline. Figure 1 shows one rule of the GL (about inline tutorials), the categorisation in the priorities and a comment given to this rule.

4 Future Plans and Development

The future plan is to have a useful and usable set of GL for game development like the W3C/WAI guidelines [4] for web pages. These GL are a standard for developing accessible web pages. One of the reasons for the great success of these GL is the possibility to test web pages if they are fulfilling the GL with programs like Bobby [5].

The future development of the games GL has two main phases, in phase one are short range goals for the GL itself, in phase two are goals for the adjustment of the GL and the distribution.

To get more in the detail, the main issues in phase one are:

- Start co-operation or to intensify co-operation with other organisations and companies working in the area of games accessibility.
- More explanation and, if possible, add best practice examples and code samples to the rules.
- Implement life cycles and visioning in the GL.
- Stronger integration of AT in the GL, specially adding an input device section in the GL.
- Ideas for dissemination and marketing.

Not all of the issues of phase two will practical, but we will try to implement as much as possible. Here are the main issues of phase two:

- Introduce a certification for games fulfilling the GL (or parts of it).
- Implement a testing tool for games, testing the observe of the GL.
- Split the GL up for different platforms, e.g. PC, Xbox, mobile devices, …
- Split the GL up for different user groups, e.g. GL for small children, GL for medical testing games, …
- Add a search function to the GL, so that a game developer can e.g. search for all relevant rules for blind people on PCs or deaf people on mobile devices.
- Add the cultural aspect to the GL.

A very important issue is, as already mentioned in the listing above, the integration of AT in the guidelines. We decide to make an additionally document to the guidelines, in which the different kinds of AT's are listed with the possibility to add devices and descriptions to every kind of AT. The main idea of this additionally document is to give game developers an overview of all the AT's being available and to give them hints and ideas how to integrate these technologies in there games.

The future goal is to have dynamic GL, which are continuously upgraded with new ideas and to be adapt to new requirements. If these GL are used for developing games, not only disabled people benefit, also the “normal” gamer benefit form an easier to understand and more configurable computer game.

Furthermore these GL will be part of an AGA (active games accessibility) framework, where games developers will have, beside these GL, a set of tools, code samples and functions for an easier development of accessible games.

The other parts of this AGA framework is still in the research and development phase, the next step will be a descriptive language. This language will make the connection between the game engine and all input and output devices, used in a game, with the option of a kind of free configuration between the different devices and the possibility to use more devices simultaneously. In addition, this language can be used to adapt already developed games to make them accessible.

References

1. http://www.igda.org/accessibility/IGDA_Accessibility_WhitePaper. pdf
2. http://www.medialt.no/rapport/entertainment_guidelines/index.htm
3. http://gameaccess.medialt.no/guide.php
4. http://www.w3.org/TR/WCAG10/wai-pageauth.html
5. http://webxact.watchfire.com/

A Classification, Based on ICF, for Modelling Human Computer Interaction

M. Billi, L. Burzagli, P.L. Emiliani, F. Gabbanini, and P. Graziani

Consiglio Nazionale delle Ricerche Istituto di Fisica Applicata "Nello Carrara"
Via Madonna del Piano, 10 – Sesto Fiorentino, I-50019, Italy
{m.billi, l.burzagli, p.l.emiliani, f.gabbanini, p.graziani}@ifac.cnr.it

Abstract. The present paper describes a study on the use of the International Classification of Functioning Disability and Health (ICF) of the World Health Organization (WHO) for modelling the interaction between humans and devices within the context of the Information Society. An extension of the human abilities section of ICF and an ICF-like classification for ICT interactions and the environment are proposed.

1 Introduction

An exact identification of user disabilities and abilities and the corresponding possibility of coping with the interactions necessary for accessing new devices, services and applications are of paramount importance, both for the activity of professionals in rehabilitation technology and for providers aiming to deploy services that take into account the needs of all users, including those with disabilities.

ICT traditionally deals with user capabilities by grouping users into categories: the blind, the visually impaired, the deaf, etc. This approach has been increasingly criticized by associations of people with disabilities, because specification of the categories is only approximate. Moreover, the classification of device features and functionalities has been considered mainly from the perspective of service providers who need to know the general constraints for a correct delivery of services [1], rather than the problems faced by users interacting with said services.

The aim of the present paper is to direct the discussion towards an identification of a formally correct methodology for identifying the capabilities of interaction with ICT equipment and services/applications which will take into account user abilities/disabilities, the context of use, and the interaction modalities offered by state-of-the-art devices. As a first step in the development of this methodology, an extension of WHO's ICF (International Classification of Functioning Disability and Health) [2] is proposed, which is aimed at making possible an easy matching between user capabilities and the interaction features of different devices and services/applications, and which also includes any constraints imposed by the context of use.

Even if biased at present from a medical point of view, the ICF classification appears to be a suitable choice, thanks to the following features: ICF is a widely-accepted standard; it was conceived as an open classification that can be extended to describe additional characteristics (e.g. users and contexts of use); and it provides for a representation with codes, thus lending itself to a mechanisation at different levels of the reasoning relative to users and ways for compensating for possible interaction difficulties.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 407 – 414, 2006.

Lastly, it is necessary to emphasize the important change in perspective promoted by the ICF document. In fact, it provides an approach for classifying abilities and disabilities (and not only disabilities) as an interactive and evolutionary process in which an individual's functioning in a specific domain is a complex relationship between his/her health condition and contextual factors.

2 The ICF Classification

ICF is a classification of health and health-related domains that describes changes in body function and structure. It has two parts, each with two components*:* part 1 deals with Functioning and Disability, and consists of two components: Body Functions and Structures and Activities and Participation; part 2 deals with Contextual Factors, and consists of two components: Environmental Factors and Personal Factors.

Each component can be expressed in both *positive* (abilities) and *negative* (disabilities) terms. It consists of various domains, and categories and subcategorie*s* are defined within each domain, as in the following example:

```
FUNCTIONING AND DISABILITY (part)
  ACTIVITIES AND PARTICIPATION (component)
    Communication (domain)
      Communicating - receiving (d310-d329) (categories)
        d315 Communicating with - receiving - nonverbal mes-
             sages (category)
          d3150 Communicating with - receiving - body ges-
                tures:. (subcategory)
```

The health and health-related states of an individual may be described by selecting the appropriate category code or codes and then adding qualifiers. These are numerical codes that specify the extent or magnitude of the ability or disability in that category, or the extent to which an environmental factor is a facilitator or barrier, according to the following scheme (where xxx stands for the second-level domain number):

xxx.0 NO problem (none, absent, negligible,…) 0–4 %
xxx.1 MILD problem (slight, low,…) 5–24 %
xxx.2 MODERATE problem (medium, fair,...) 25–49 %
xxx.3 SEVERE problem (high, extreme, …) 50–95 %
xxx.4 COMPLETE problem (total,…) 96–100 %
xxx.8 not specified
xxx.9 not applicable

If necessary, time-related conditions can also be specified and quantified.

The use of ICF to describe user needs related to ICT activities, such as technology-mediated communication and access to information, can take advantage of a number of features in both its qualitative and quantitative representations:

- the identification and quantification of the degree of ability or disability possible with ICF coding can be highly detailed, even to the point of describing individual characteristics;
- the need for using facilitators can be stated by means of the differentiation between capacity (related to personal abilities) and performance (related to personal conditions and context);

- the possibility of quantifying abilities to perform specific actions can be crucial to identifying the most suitable facilitator and service-delivery modality;
- the impact of the context as environment-related conditions can be specified and quantified in a highly detailed way.

3 Proposed ICF Extensions

In order to identify user capabilities in connection with having access to information and with interpersonal communication in the information society, it is necessary to describe the users themselves, the interaction characteristics of the technology employed, and the context of use (including usage of the environment, personal factors, and possible modifications of the interaction technology). Here as follows modifications to the ICF classification of users and the necessary extensions of the classification of the technology used are proposed. The need for modifications within the context classification is outlined, but is not developed in detail.

3.1 ICF Coding of User Interaction Capabilities

ICT needs a number of different ways of interaction, depending on the available devices, the actual environment, and the context of use. Ability in required interactions is connected not only to personal conditions (for example, a car driver is functionally blind and motor-disabled as far as interaction with a navigator is concerned). Therefore, the proposed classification, which has been hypothesised in its initial status within the MAIS project [3], focuses on the current functional status of users, changing dynamically under the influence of the environment and the situation, rather than on the static ability/disability status of the user. This enables users to specify their needs within a context of use, in addition to that of their health condition, impairment, or other limitations.

In this scenario, when describing "body functions", ICF is an appropriate tool. Instead, if there is a need to represent user actions, the heuristic ways used in ICT are specific, but do not lead to a formal classification and coding of such actions, nor to structures and qualifiers [4, 5, 6, 7]. The aim of the present work is to propose a formal classification of user actions by extending the ICF classification and by adopting the ICF method for quantifying capacity and performance, as well as by taking into account the influence of the context on performing such actions.

In a consideration of the various abilities needed to interact with a device for accessing and producing information or for communicating, three areas can be identified:

- the comprehension area[1], which groups all the actions related to receiving and understanding information or communication in every possible form;
- the expression area[2], which groups all the actions related to producing information or communication in every possible form;
- the skills area, which groups all the actions needed to use the chosen device.

[1, 2] It should be noted that comprehension and expression capabilities are identified as purely conceptual and sensorial actions, distinct from the skills needed to make it possible to perceive or express something by means of technological devices.

By comparing the specified areas with ICF, it can be seen that the first two are very similar to the "Communicating – receiving" and "Communicating – producing" categories, with some minor differences (cfr. ICF Activities - Communicating – receiving (d310 – d329) and Communicating – producing (d330 – d349)).

On the contrary, as far as skills are concerned, the ICF classification does not contain the richness of detail that is needed in order to completely describe the actions to be performed when interacting with a system for accessing and producing information or communication.

The proposed solution is to define a completely new domain with respect to ICF, named "Skills" (see appendix A), and to assign to it the character "k" as code prefix [3]. In defining the new domain, the interaction techniques common to current devices have been taken into account. A small role has been assigned to special techniques, such as, for example, virtual reality, which is employed in limited and very specialized fields.

3.2 ICF Coding of Interaction Characteristics of Devices

Several specifications have been proposed for describing devices, such as Composite Capabilities/Preference Profiles (CC/PP), Media Features, Media Queries, User Agent Profile Base Vocabulary (UAProof), etc.. Irrespectively of the richness of the vocabularies adopted, their aim seems to be limited to representing characteristics of interest for service operators, rather than those of users. This approach is, of course, relevant in designing services, but the present study emphasises the need for a well-defined way of matching user needs with the interaction characteristics of devices.

Therefore, a classification of device functionalities is proposed, in parallel with the one depicted in the preceding section on user characteristics (see appendix B). By using the philosophy underlying the ICF coding, a classification has been defined for determining the "health status of a device" by comparing it with an ideally capable device, that features all the interaction capabilities offered by the present status of the technology. Consequently, by using a simple coding set, it will be possible for the actual device functionalities to be concisely classified, in order to define the possible range of communication and interaction when a user employs a given device.

As in the previous section, a division into three domains has been adopted: communication reception, communication production, and interactions[4].

Voice recognition does not appear among the audio characteristics of a device, since, at the moment, it is not considered as a possible means of communication. However, it is mentioned twice in the interaction section, since voice recognition can be used both to input text and to perform pointing actions.

3.3 ICF Classification and Contexts of Use

The context of use is crucial in the classification presented. In fact, user abilities and device capabilities are not considered as static, but as changing dynamically depending on the changes in the context of use.

---

[3] In the following user-needs classification, cognitive disability has not been taken into account. This is because it is unique compared to other disabilities, and its classification lacks a clear correspondence with interaction tasks, as significantly expressed by [8].

[4] Communication reception should be understood as a communication received by a user and, therefore, produced by a device.

The definition of context varies widely in the literature, ranging from a set of environmental characteristics [9] to a philosophical view of context as being a component of the interaction itself [10]. In this study, context is defined as the set of factors that can influence the interaction between users and devices. It can be divided into three domains: environmental factors, situation conditions, and device extension, integration or modification.

Environmental factors, such as light condition, noise, etc., influence the capability of perception of screens and audio, which is equivalent to diminishing the capabilities of the devices. Situation conditions, such as travelling (by plane, by train, by bus), driving, walking, etc., influence both user abilities and device capability. Device extension, integration or modification, such as the use of extended or reduced keyboards, screen readers, mouse emulators, etc., alters device capabilities.

A complete study of the relevant context features and their influence on user abilities and device capabilities is still in progress.

4 Examples of Use of the Classification

Two possible utilisations of the proposed classification are suggested in order to show the potential of the classification to offer a basis for structuring, in a consistent and easily comparable way, our knowledge of the interaction between users and devices, by taking into account the fundamental role of the context of use.

4.1 Use in Rehabilitation

The proposed classification could be an important aid in structuring the functional assessment of disabled users by rehabilitation professionals, and in identifying the best solution for interaction with ICT equipment and services/applications. Such a classification can be developed in a consistent and systematic way so that to be comparable throughout all the rehabilitation centres which adopt it.

Rehabilitation operators could match the results of the user assessment with the interaction characteristics of available devices by taking into account the influence of the context. The said matching could be backed up by computer procedures that structure the assessment dialogue and provide support in the decision-making process, generating a range of solutions ordered on a scale of predicted effectiveness.

4.2 Use in Modelling Human-Computer Interaction

Within the more general scenario of HCI, the use of classifications could help in dynamically determining the interaction possibilities available when a user accesses a service via a specific device in a varying context of use.

The characteristics of the context of use, as previously defined, can determine the actual user-interaction capabilities, which are determined both by the abilities that the user can actually employ under the environmental and situational constraints and by the characteristics of the available interaction device. Similarly the usable features of the device are correspondingly determined by the environment and by the availability of facilitating modifications. These pieces of information can give the service provider a real-time hint on how best to deliver the requested service so as to maximise service effectiveness under all possible conditions of use.

5 Conclusions

The aim of this paper is to initiate a discussion on the need for and possibility of facilitating the matching-up of user abilities/disabilities with the interaction capabilities of state-of-the-art devices. To this end, the development of a threefold classification (user skills, device characteristics, environment status) is proposed, based on the ICF classification and its innovative philosophy, according to which it can be extended while still maintaining its strength and consistency.

The possibility of classifying users, the interaction features of devices, and contexts of use with the quantification deriving from the qualifiers inherent in the ICF classification methodology has been shown.

The coding and numerical qualifiers suggest the possibility of a mechanisation of the reasoning relative to the matching-up of users with the interaction capabilities of devices.

References

1. User Agent Profile 1.1 - Candidate Version 12-December-2002 Open Mobile Alliance OMA-WAP-UAProf-v1_1-20021212-c2.
2. ICF International Classification of Functioning, Disability and Health. World Health Organization, Geneva (2001)
3. Graziani, P., Billi, M., Burzagli, L., Gabbanini, F., Palchetti, E., Bertini, E., Kimani, S., Sbattella, L., Barbieri, T., Bianchi, A., Batini, C.: Deliverable 7.3.1 Definition of User Typologies. Multichannel Adaptive Information Systems a FIRB Project.
4. Kobsa, A.: Generic User Modeling Systems. In User modeling and User-Adapted Interaction 11. Kluwer Academic Publishers. The Netherlands (2001) 49-63
5. Fisher, G.: User Modeling in Human-Computer Interaction. In User modeling and User-Adapted Interaction 11. Kluwer Academic Publishers. The Netherlands (2001) 65-86
6. Kay, J.: User Modeling for Adaptation. In Sthephanidis, C., User Interfaces for All. Lawrence Erlbaum Associates Publishers. London (2001) 271-294
7. IST 20656-0 PALIO Del 01 (2202) - "User needs and characteristics and contexts of use"
8. Vanderheiden, G.: Fundamental and Priority for Design of Information and Telecommunication Technologies. In Presiser, W.F.E., Ostroff, E., Universal Design Handbook. McGraw-Hill New York (2001) 65.3-65.15
9. Schilit, B.N., Theimer, M.M.: Disseminating active map information to mobile hosts. Network, IEEE Volume 8, Issue 5, Sept.-Oct. 1994 Page(s):22 - 32.
10. Dourish, P. : What We Talk About When We Talk About Context. Personal and Ubiquitous Computing, 8(1), 19-30. Springer-Verlag London Ltd February 2004

Appendix A: Classification of Skills

- k0 Keys/buttons control
  - k00 Functional
    - k000 To press a single key/button
    - k001 To press multiple keys/buttons
    - k002 To control a key/button pressing over time
  - k01 Format
    - k010 Reduced
    - k011 Extended
    - k012 Virtual (on screen) key/button
    - k013 Projected or holographic key/button
- k1 Write capable
  - k10 Using a keyboard
    - k100 QWERTY
    - k101 Keypad
      - k1010 Using disambiguation (e.g. T9)
      - k1011 Without disambiguation
    - k102 Chord
  - k11 By hand on a screen
  - k12 Articulating voice to input text
- k2 Pointing
  - k20 To move the pointer
    - k200 By fine movement with hand-eyes coordination (e.g. mouse, touch pad)
    - k201 By rough movement with hand-eyes coordination (e.g. joystick, trackball)
    - k202 By to exert a pressure in coordination with eyes (e.g. using the keypad to move the pointer, or using 5 buttons to make choices)
- k3 Clicking
  - k30 To make a click
  - k31 To make a double click
  - k32 To make a right click
  - k33 To drag
- k4 Making gestures
  - k40 To rotate the mouse wheel
  - k41 Touch screen (direct pointing)
    - k410 To touch a portion of the screen with a finger
    - k411 To touch a portion of the screen with a pen
    - k412 To make gestures on a portion of the screen with a finger
    - k413 To make gestures on a portion of the screen with a pen
  - k42 Touch tablet (absolute indirect pointing)
    - k420 To touch a tablet with a finger
    - k421 To touch a tablet with a device
    - k422 To make gestures on a tablet with a finger
    - k423 To make gestures on a tablet with a device
  - k43 To move the phalanges (e.g. to use a data glove)
- k5 Articulating voice to interact

Appendix B: Classification of Interaction Characteristics of Devices

- x1 *Communicating - receiving*
  - x10 Screen
    - x100 Screen size
    - x101 Colour capable
    - x102 Text capable
    - x103 High contrast
    - x104 Graphics/Image capable
    - x105 Animation/Video capable
    - x106 3D capable
  - x11 Audio
    - x110 Sound capable
    - x111 Playback capable
    - x112 Synthetic speech capable
    - x111 Conversation capable
  - x12 Braille capable
  - x13 Tactile feedback capable
- x2 *Communicating - producing*
  - x20 Text capable
  - x21 Colour capable
  - x22 Producing Graphics
    - x220 Graphics/Image capable
    - x221 Animation/Video capable
  - x23 Producing audio
    - x230 Sound capable
    - x231 Record capable
    - x232 Conversation capable
- x3 *Interactions*
  - x30 Typing
    - x300 Keyboard type
      - x3000 QWERTY
      - x3001 Keypad
      - x3002 Chord
    - x301 Keyboard format
      - x3011 Reduced
      - x3012 Extended
      - x3013 Virtual (on screen)
      - x3014 Projected or holographic
    - x302 Keyboard Options
      - x3020 Serialization
      - x3021 Rebound control
      - x3022 Key depression time control
      - x3023 Disambiguating
      - x3024 Time dependant
  - x31 Pointing & clicking
    - x310 Mouse like
    - x311 Touch pad
    - x312 Joystick like
    - x313 Buttons
    - x314 Touch screen
    - x315 Touch screen, requiring pen
    - x316 Touch tablet
    - x317 Touch tablet, requiring device
  - x32 Gesture
    - x320 Mouse wheel
    - x321 Gestures capable on screen
    - x322 Gestures capable with a pen on screen
    - x323 Gestures capable on tablet
    - x324 Gestures capable with a device on tablet
    - x325 Handwriting on screen
    - x326 Handwriting on tablet
    - x327 Manipulating 3D objects
  - x33 Voice recognition
    - x330 Articulate to input text
    - x331 Articulate to interact

Evaluation of Reaction Forces During Human Computer Interaction for Optimization and Development - A Pilot Research

Stefan Mina, Lisa Ehrenstrasser, Robert Lurf, and Barbara Prazak

ARC Seibersdorf research GmbH, Rehabilitation and Inclusion Research Group,
A-2700 Wr. Neustadt, Austria
stefan.mina@arcsmed.at
http://www.arcsmed.at

Abstract. The described work focuses on the analysis how users interact with specific human machine interfaces and how shape and tactile properties influence this interaction. Results gained by such investigations should help to obtain better knowledge of human computer interaction in general and about possible influences of diseases.

Based on this knowledge it should be possible to fit and optimize human machine interfaces in a way that is not only dependent on the professionist performing this task. Also for the development of new device this knowledge and reliable data is essential.

1 Introduction

The Selection of input devices in the field of rehabilitation technology is very often only based on trial and error or on the personal experience of the person who selects the device.

Even in the development and design process of new human machine interfaces developers and designers only rely on currently available systems or do not consider user interaction and user needs at all.

Usability research is commonly found in market oriented investigations of companies or institutions who want to test their products [1]. There can be found also international standards for human computer interfaces and usability like described by Bevan [2]. The application of procedures and apparatus used in this field of research can only rarely be found in the field of rehabilitation technology and in development of human computer interfaces for people with disabilities. The introduction of such technologies could significantly improve the knowledge about human machine interaction, the influence of specific diseases and the influence of device design so that optimal solutions for people of the intended target group can be found. Currently an increasing number of such research can be observed and the results seem to be promising like some examples show [3,4].

Aim of the described work is to analyze how users interact with human machine interfaces and how this interaction can be influenced by design issues. In the current stage of the project a focus is set on the analysis of reaction forces,

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 415–420, 2006.

other important features like motion analysis will be part of future work. With an extended test battery the process of design of new human machine interfaces for people with manual impairment and the selection of interfaces for this target group can be significantly improved.

2 Methods

One component in the analysis of user interaction are the forces applied by the user onto the input device. These reaction forces reflect on the one hand the abilities of the user depending from their disease and on the other hand the usability of the input device itself. To measure the reaction forces a multiaxis force plate, where different input devices can be mounted, was developed (shown in figure 1). This measurement device is based on strain gage technology and has the capability to measure forces in four axes. Data acquisition is performed with a strain gage amplifier (IOtech DBK43A), a data acquisition device (IOtech DAQbook 2000/E) and a notebook. For further investigations including electrophysiology, especially electromyography (EMG) an eight channel biosignal amplifier system (two Grass QP511 quad AC-amps) was connected to the data acquisition device. Post processing of the acquired data (sample rate 256 Samples/second) was performed with Matlab and SPSS.

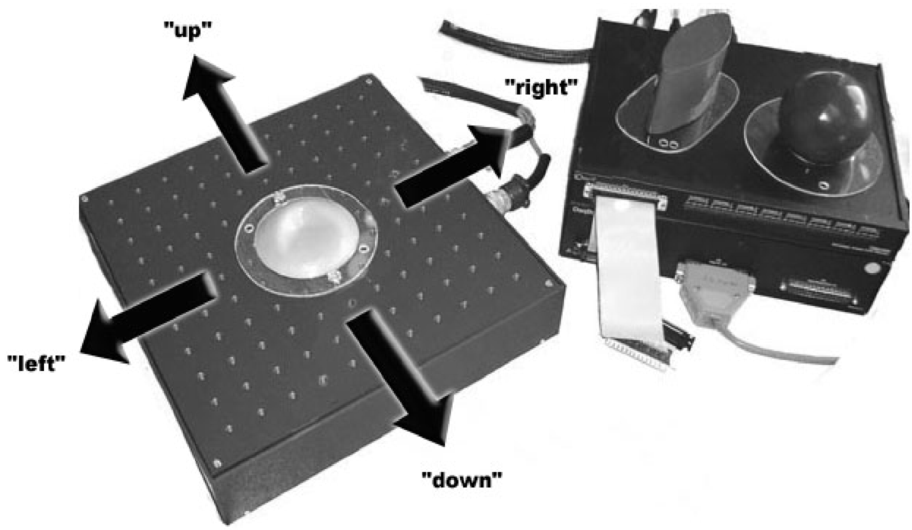


Fig. 1. Force Measurement Plate

Similar force measurement plates can usually be found in biomechanical research or in industrial applications. Biomechanical research for example focuses on the analysis of human gait and the influence of different factors by using force plates [5].

Able-bodied users carried out all experiments, since the intention was only to proof the methodology. Of course it would be necessary to involve users of the intended target group (people with physical disabilities) but this would required a well planned study with homogenous test groups since differences between specific diseases may be expected and are subject to future studies based on the system and the methodology described in this work.

2.1 Optimization and Fitting of Existing Human Machine Interfaces

The first part of the measurements focuses on the analysis of interaction with common input devices like joysticks.

Therefore a standard 2-axis joystick was mounted on the force measurement plate and a sample of users had to perform a specific task. The task consisted of a simple sequence (up, down, left, right), which had to be repeated 3 times by the user. To remove unwanted components by pacing, the user was only verbally instructed about the task before beginning and the sequence was neither visually nor verbally presented to the user during performing the experiment.

Similar investigations with pointing devices (headmouse, mouth operated mouse etc.) were already carried out by the author [6] where the performance of these devices was tested. Results of the study were already used for product optimization of the Lifetool IntegraMouse, which is a mouth operated mouse.

2.2 Development of Novel Human Machine Interfaces

The main design question is how surfaces and shapes manipulate the users force, even more give either positive or negative magnification. To intensify the different force output made from the user in order to find out what is increasing and what is decreasing, variable basic shapes have been built. These basic shapes are split in three experiment areas:

Basic shape A: for the whole hand (red forms)
Basic shape B: for the palm (blue forms)
Basic shape C: for fingers (either index finger only or index finger, middle finger and ring finger together, yellow forms)

Basic shapes A, B and C are influenced by the grasp reflex and the resting hand pose (without muscular tension). The grasp reflex stops at young age, still it influences the intuitive grip - grasping for something with a lot of force.

'Grasp reflex: A type of reflex in which the fingers flex in a clutching motion in response to stroking of the palm of the hand. It is normal in young infants from birth to about four month of age [7].'

The haptic sense as a tactile force together with surface sensibility gives the background for a future user experiment, which would compare different textures, finishes and stiffness of the materials.

Tasks, which had to be performed where similar to them described in 2.1. The user had to push or pull the input device in all four directions (left, right, up, down) with a maximum affordable force. Due to the fact that only able-bodied

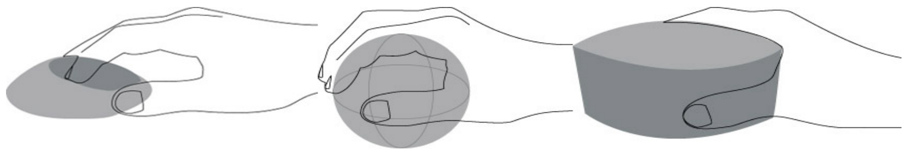


Fig. 2. Basic Shapes (A,B and C)

users performed the experiment it can be expected that the users maximum force is almost equivalent in all four directions. When dealing with impaired users the affordable forces may be influenced by the pathology. Such differences can be taken into account when fitting the input device to a specific user by compensating the force differences by an appropriate signal processing or mechanical design.

3 Results

Data from the experiments according to 2.1 from 15 test persons were evaluated using statistical methods. Before applying these methods the time discrete data-set of each individual containing the whole test sequence was separated into single force data-sets for each joystick position using Matlab.

After this preprocessing it was possible to calculate the mean and standard deviation of each joystick position (up, down, left, right) with an additional removal of bias forces (derived from the values with the joystick in center position). The histograms in fig. 3 are showing an example evaluation of the mean X,Y and Z force when pushing the joystick to the "up direction". Please note that in all evaluations a biomechanical kinematic naming of the force directions according to Winter [8] is used.

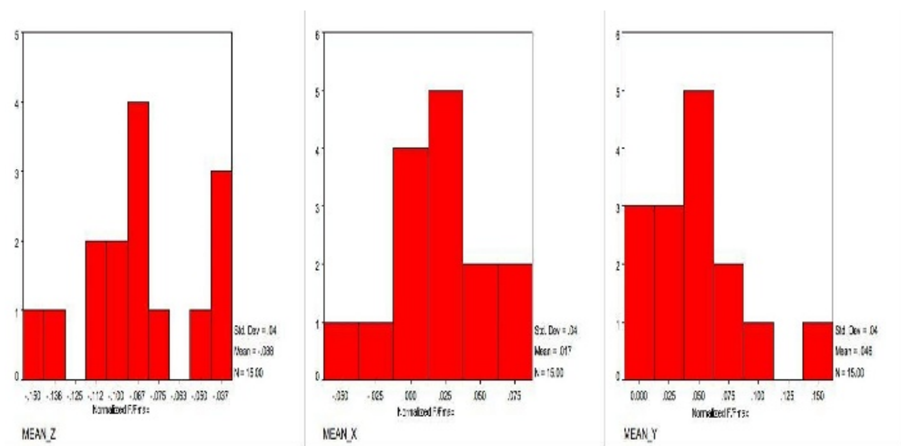


Fig. 3. Histogram of Forces

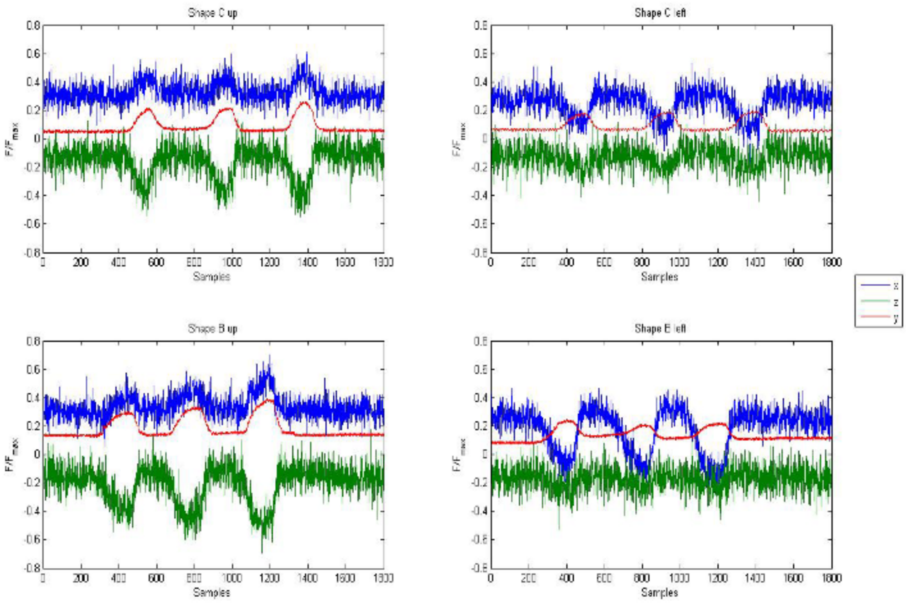


Fig. 4. Comparison of reaction forces using different shapes

The evaluation shows the distribution of the mean force of all 15 test users which can be useful for the design of a joystick to set the correct force thresholds (forces are scaled to maximum force).

The evaluation of different shapes for novel input devices shows interesting results. Figure 4 shows 3-axis force plots of an experiment using two different shapes (basic shape B and C). The user had to push the shape three times upwards and three times leftwards. The two left plots in 4 show that there is almost no difference in the force verctor between shape B and C when pushing upwards (-Z direction). When pushing to the left (X direction) it could be noticed that shape B (lower plot) provides a better force transduction than shape C which is not symmetric. Of course it appears to be obvious that a symmetric shape provides a better force transduction than an assymetric one but it is more difficult to estimate the behaviour with complex shapes, if the user is influenced by a disease.

4 Discussion

Based on the results and the experiences made the introduction of standardized measurements as decision support for the selection of input devices and for the development of novel devices seems to be crucial.

Since in some countries technical aids for people with disabilities are not or only partially paid by health insurances and the selection is not properly done many users have to use suboptimal input devices and other equipment. A selection tool based on standardized methods and basic knowledge could be beneficial for the user to get an optimized solution and improve the quality of life.

However not only the user itself benefits, the designers and product developers process for redesign and new design can be inclusivly extended in their threedimensional output in a more adequate userorientated product.

References

1. Maguire M, Bevan N (2002) User requirement analysis: A review of supporting methods. In: Proceedings of IFIP 17th World Computer Congress, Montreal, Canada, pp. 133-148, Kluwer Academic Publishers
2. Bevan N (2001) International Standards for HCI and Usability. International Journal of Human Computer Studies, **55(4)**, 533-552
3. Rodriguez MG, Perez Perez JR, Paule Ruiz MP (2004) Designing User Interfaces Tailored to the Current Users Requirements in Real Time. In: K. Miesenberger et al. (Eds.) ICCHP 2004, LNCS 3188, pp. 69-75, Springer Verlag Berlin Heidelberg
4. Simsik D,Galajdova A (2005) Accessible information and quality of human-computer interaction. In: A. Pruski and H. Knops (Eds.) Assistive Technology: From Virtuality to Reality, pp.718-722, IOS Press Amsterdam
5. Elftman H (1939) Forces and Energy Changes in the Leg during Walking. Am. J. Physiol. **125**, 339-356
6. Mina S, Hochgatterer A, Prazak B, Hanke S, Schreier G (2004) Multi-parameter Data Acquisition on Interface Devices for the Development of Adaptive Input Systems. In: K. Miesenberger et al. (Eds.) ICCHP 2004, LNCS 3188, pp. 37-44, Springer Verlag Berlin Heidelberg
7. Colman, Andrew M.; Grasp reflex p. 315; Oxford, Dictionary of Psychology 2001 by Oxford University Press
8. Winter DA (2005) Kinematics: Kinematic Conventions. In: Winter DA (Ed.) Biomechanics and Motor Control of Human Movement, 3rd Edition, pp. 14-15, John Wiley and Sons New Jersey

A User-Orientation Evaluation Framework: Assessing Accessibility Throughout the User Experience Lifecycle

Alexandros Mourouzis[1], Margherita Antona[1],
Evagelos Boutsakis[1], and Constantine Stephanidis[1, 2]

[1] Foundation for Research and Technology – Hellas (FORTH),
Institute of Computer Science, 71300 Heraklion, Greece
{mourouzi, antona, boutsak,cs}@ics.forth,gr
[2] University of Crete, Department of Computer Science, 71 409 Heraklion, Greece

Abstract. Today, as the users and contexts of use of Information and Communication Technology (ICT) become more diverse, there is a significant need to understand all the factors that may affect the quality of the user-experience, and to measure them systematically. This paper proposes an evaluation framework for specifying and measuring the *user-orientation* of interactive products. The term "user-orientation" refers to the extent to which target users (will) find the product acceptable at all phases of the user experience lifecycle. The proposed framework incorporates accessibility as a basic determinant of acceptability and long-term adoption. It can be employed effectively in the evaluation of systems that are aimed to be accessible and usable by diverse users (e.g., public systems) or by people with disability.

1 Introduction

In order to develop better methods for designing new computer technology, testing designs and predicting how users will respond to new products, it is important to understand why a person may fail or resist using an interactive product. Since the (target) user population of computer technology became more diverse, researchers and practitioners realised that *utility* and *usability* alone, although salient acceptance factors, do not necessarily imply that take-up rates of a system will reach their full potential. For instance, in designing and delivering information technology for the public, *accessibility* needs to be guaranteed, and interfaces are required to cope with a variety of users, including people with different cultural, educational, training, and employment background, novice or experienced users, the very young and the elderly, and people with different types of disability. In fact, accessibility is considered as a prerequisite for usability, since there cannot be optimal interaction if there is no possibility of interaction in the first place [1].

As a result, traditional usability-oriented design and evaluation approaches are not sufficient for delivering high-quality systems that are able to accommodate the needs of diverse user populations with different abilities, skills, requirements and preferences, and to be used in dynamic contexts of use (tasks, equipment and environmental conditions). An analysis of the involved sources of variability reveals the necessity of developing and adopting more comprehensive approaches, capable of specifying and

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 421 – 428, 2006.

measuring the *user-orientation* of products and their interfaces. In this paper, the term "user-orientation" refers to the extent to which target users (will) find the product acceptable at any phase of their 'interaction' with it, i.e., throughout the entire user experience lifecycle [2].

This paper defines a *user-orientation evaluation framework*, which incorporates accessibility as a basic determinant of system acceptability, and describes how such a framework can be employed effectively in the evaluation of systems and their interfaces, including systems that are aimed to be accessible and usable by diverse users (e.g., public systems) or systems developed for people with disability.

2 Research and Methodological Approach

Common models of technology acceptance have their roots in a number of diverse theoretical perspectives. One of the most popular is that of Innovation Diffusion Theory [3], which seeks to identify significant perceived characteristics of technology which may impede its adoption by users. On the other side, in social psychological research, theorists seek to identify determinants of behaviour within the individual rather than the technology. For instance, the Theory of Reasoned Action (TRA) [4] has been used to investigate how user beliefs and attitudes are related to individual intentions to perform. TRA provides a complete rationale for the flow of causality from external stimuli (such as system design features) through user perceptions to attitudes about the technology, and finally to actual usage behaviour.

Therefore, for the definition of a comprehensive evaluation framework ,one needs to determine, on the one hand, whether a user *can* use a product - subject to the *characteristics of the user* and the *context of use*, and on the other hand, whether a user *will* use the product - subject to the *user's behavioural situation* (see Fig. 1). In terms of user characteristics, key personal differences may be characterized as physiological, psychological or socio-cultural [5] and may involve variances in gender, physical and cognitive abilities, language, culture, experience, background, etc. The context of use involves diversity in terms of user tasks, equipment (at the user site), and in social and environmental conditions [6]. Finally, in terms of the user's behavioural situation, the Technology Acceptance Model [7] derived from TRA identifies two salient factors: *perceived usefulness* and *perceived ease of use*. Perceived usefulness can be considered as the degree to which a potential user believes that using a particular system would match individual goals, and perceived ease of use as the degree to which a potential user believes that accessing and using a particular system would be free of effort. More recent approaches have identified additional constructs, such as, for example, *perceived risk* (PR), i.e., the degree of uncertainty regarding possible negative consequences of using a product. In [8] the authors propose that PR comprises the facets of performance, financial, time, psychological, social, privacy and overall risk. Furthermore, when measuring user satisfaction as a key determinant of user acceptance, the fact that satisfaction is heavily influenced by *expectations* should be taken into account. In other words, perceptions of product quality stem from a comparison of what customers feel a product should offer (i.e., their expectations) with the way the functions are actually delivered. A number of factors influence user expectations, e.g., previous experience, personal needs, implicit service communication, values and beliefs, views about the provider, explicit service communication, and word-of-mouth communication [9].

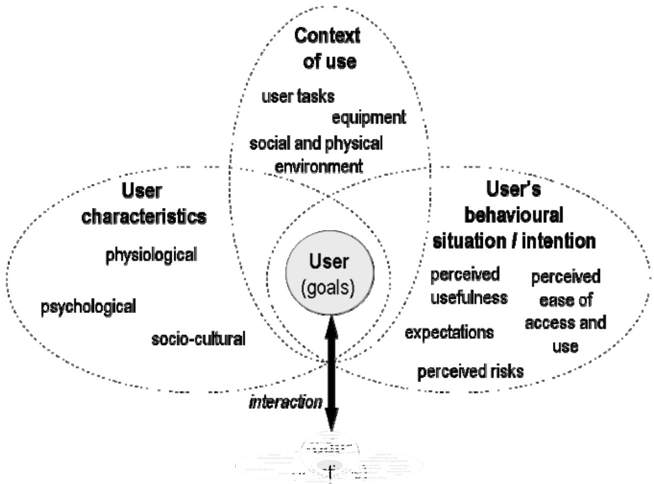


Fig. 1. *Conditions of use*: user's characteristics, context of use and behavioural situation encompass determinant acceptance factors towards and during the use of a product

3 A Comprehensive Evaluation Framework

3.1 A Systematic Perspective

In the development of the proposed approach, a number of theoretical and empirical models of technology acceptance were considered. Emphasis though has been put on the user's decision making process towards the use and adoption of a system. The main generic perspective looks at the process of discovering, identifying the need or interest to use, reaching (i.e., accessing), using, and re-using a system (see Fig.2).

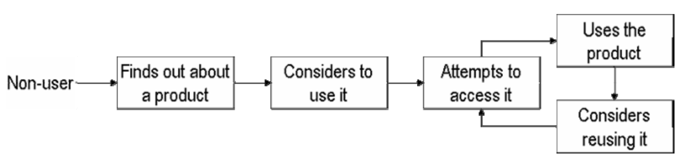


Fig. 2. An abstract model of the user experience lifecycle

Regarding the phase of product usage, and towards long term adoption, previous studies based on empirical data [10] have identified a sequence of three patterns of use: (a) *exploration in width*, which refers to the phase of the preliminary familiarisation of the user with the interactive environment of a system; (b) *occasional (long* *term) use*, which refers to the phase of common use of a system; and (c) *exploration in depth*, which refers to specialised usage.

The proposed approach implies careful consideration of all the factors which may influence an individual at each phase of the lifecycle and concluding upon the likelihood that the individual will proceed or not to a subsequent phase. Ultimately, this allows predicting or assessing the likelihood that a non-user will eventually become a faithful product user (i.e., the evaluating acceptability of the system).

3.2 Framework Definition

User-orientation (and thus system acceptability) is measured by the extent to which:

– the product[1] is made visible to non-users (*visibility*),
– non-users are motivated to gain a personal experience of the system (*perceived usefulness & ease of use*),
– actual users find it easy and acceptable to reach the product (*availability/approachability*),
– actual users find it useful, easy and acceptable to interact with the product (*quality of interaction experience*),
– previous users are motivated to become long term users (*relationship maintainability and subjective usefulness & ease of use*),
– product users are not offered more promising and satisfying alternatives (*competitiveness*).

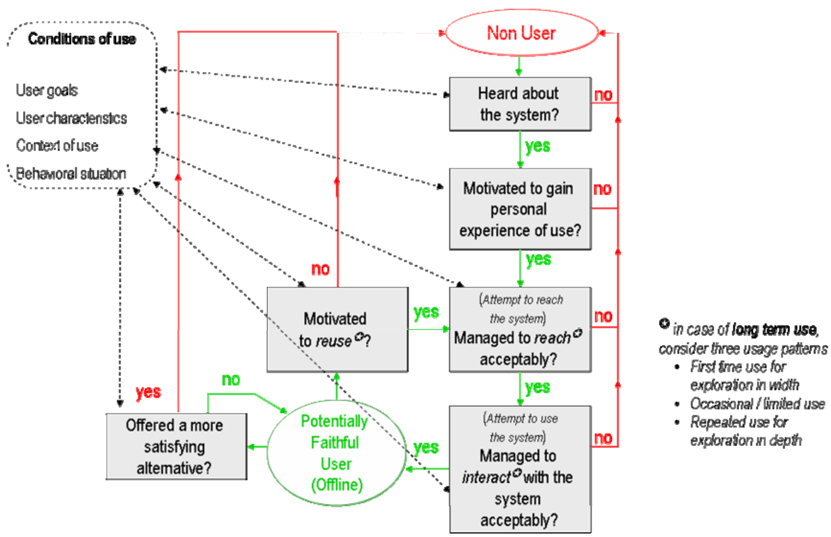


Fig. 3. Overview of the user-orientation evaluation framework

[1] Note that the product can be of various types including service, software, hardware , function or any combination of these.

Emphasis is given to the *Conditions of Use* (see Fig.3) and to the fact that the level of user-orientation depends on the specific circumstances in which the product is used, including specific *user goals*, *user characteristics* (physiological, psychological, socio-cultural), *contexts of use* (user's tasks, equipment and surrounding social and physical environment) and *user's behavioural situations* (perceived usefulness, ease of access / use and risks, and expectations).

In cases of products that are aimed to be used repeatedly and to be adopted for long term use, emphasis is also given to the fact that the level of user-orientation depends on the current phase of usage lifecycle, including *first time and novice usage* (seeking product exploration in width), *moderate usage* (seeking occasional product exploitation), and *expert usage* (seeking exploration and exploitation in depth).

3.3 User-Orientation Measurements

As it becomes apparent from the above definition of the framework, user-orientation reflects the overall product quality perceived by users as a total of the following product qualities (revised from [2]).

Visibility. Visibility refers to the degree to which a system can become known to individual[2] non-users. Obviously, the actual location of the system is a major visibility factor. Furthermore, visibility can be increased by providers through publicity strategies. Products can, however, be visible to a certain degree even if no promotion takes place, for example through easy location by means of popular web search engines. Awareness might result from unintentionally coming across a product, e.g., while surfing on the Internet. Naturally, the accessibility of the location of the system as well as of the publicity media used is a major factor of a system's visibility.

Perceived Usefulness and Ease of Use. These qualities refer to the usefulness and ease of (access and) use of the system from the viewpoint of individual non-users. These are related to the available information regarding the product and to the extent to which the product appears to be suitable with respect to the user's particular goals and needs. This also comprises a variety of tangible aspects, such as time and cost savings resulting from the product itself (rather than the way it is delivered). Finally, it can also incorporate less tangible aspects, such as personal intrinsic gratification that can be derived from the fun of, for example, participating in an attractive learning experience/training course. Other social aspects may also play a role, such as prestige and social desirability.

Availability / Approachability[3]. Availability refers to the degree to which all types of potential individual users can reach the entry point (s) of the system. Certainly, accessibility (e.g., for anyone, at any time, from anywhere) of the carrier / storage medium of a system is a major factor for its availability / approachability. At this stage, particular needs and requirements of diverse user populations, such as people with disability, are considered with regards to available 'routes' for reaching the product.

---

[2] The term "individual" user refers to individual *Conditions of Use* (user characteristics, context of use, and behavioural situations).

[3] The use of the term accessibility (in its literal sense) is avoided here in order to ensure that this is not confused with as ease of access to people with disability.

Quality of Interaction Experience. This characteristic encompasses the quality of interaction perceived by actual individual users and refers to the degree to which a system can be used to achieve useful and quality results (i.e., lead to subjective satisfaction). It can also be perceived as the total of the user-orientation of the system's functions of subjective importance to the individual user.

Relationship Maintainability (Subjective Usefulness and Ease of Use). This refers to the degree to which a good relationship with individual system users is effectively cultivated and maintained while the user is not working on the system (e.g., by means of informing the user for new functionality, content updates, changes of status, etc.). The provider may need to adopt specific strategies to ensure the sustainability of this relation. For example, the system provider may offer a richer service package in order to maximally fit the goals and needs of individual users.

Competitiveness. Competitiveness is the degree to which the system is conceived by individual users to be more appropriate for them than other available alternatives. For instance, this may be improved by informing the user about reviews of the system and market rankings.

3.4 Accessibility: A Ubiquitous Issue in the User Experience Lifecycle

Accessibility can be defined as the extent to which the sequences of input actions of a product, and the associated feedback that lead to successful product use, are possible to be performed by the user, with respect to the individual's limitations emerging from the particular conditions of use (adapted from [1]). In other words, accessibility ensures that an individual can use a product, whereas usability ensures that the individual finds it easy and satisfying to use it. Thus, further to the framework, accessibility, for instance for people with disability, is an issue penetrating all phases of the user experience lifecycle, raising questions such as "is the product (service, software, hardware, function) visible to people with disability?" or "can the entry point of the system be reached by people with disability?".

4 Employing the Framework for User Interface Evaluations

The proposed framework can also be employed to specify and measure the user-orientation of any system part (e.g., interaction devices and peripherals), including individual functions of a system and their corresponding user interfaces (UIs). In other words, when moving deeper into the evaluation of subsystems and system functions, the framework can be iteratively applied to each corresponding UI component (both physical and virtual).

Indicatively, assume a function "view incoming emails" provided through an icon in a Web mail application. The function's visibility is decreased, for example, if the icon is placed at an inappropriate location in design, or increased if information regarding the existence of this function is provided through a help module. Perceived usefulness and ease of use (i.e., prior actually using it) is decreased if the icon does not reflect the utility of the function, or increased if a help module offers a narrative description of the function's utility and guidance on using it. Reachability / approachability is decreased if the icon is not accessible (e.g., for blind users) or if it is hard to reach for motor-impaired users who navigate through tabbing. Quality of interaction starts from assessing the behaviour of the icon itself (e.g., is appropriate feedback provided while the icon is being pressed) and continues with assessing each component of the triggered dialogue. Relationship maintainability (i.e., while the user is not logged in the system) is increased, for example, through sms notifications of new incoming mails. Finally, competitiveness can be improved by offering the user information about the performance (success rates etc.) of this function.

In general, in assessing a user interface and in order to claim high levels of overall user-orientation, (a) each function needs to be highly user-oriented individually, and at the same time (b) an analogy needs to be achieved between the importance of each function to the user and the corresponding levels of user-orientation of each individual function (see Table 1).

Table 1. User-orientation evaluation of a user interface consisting of three functions

Function	Importance to the user	Function's user orientation
Function 1	High	X (*must be higher or equal to Y*)
Function 2	Moderate	Y (*must be higher or equal to Z*)
Function 3	Low	Z (*must be at least acceptable*)

5 Conclusions and Future Work

The presented framework[4] provides a holistic approach for expert-based or user-based evaluations of interactive systems, specifying at an abstract level all factors that may influence the quality of the user experience with a given system. In the context of user studies, the framework can be employed for setting goals, measurements and data collection instruments for a test, and for preparing test materials such user questionnaires, task scenarios, etc. The framework may also be employed in conducting walk-through evaluations similarly to traditional cognitive walkthroughs, which "*use an explicitly detailed process to simulate a user's problem-solving process at each step in the human-computer dialogue, checking to see if the simulated user's goals and memory for actions can be assumed to lead to the next correct action*" [11]. The main difference is that, further to the framework, the aim is to simulate a user's reasoned action process at each step in the human-computer dialogue, checking to see if the simulated user's beliefs, external stimuli (such as system design features) and intentions to perform can be assumed as a prerequisite to lead to the next interaction step. To this end, this method can be considered as a *behavioural walkthrough method*, rather than a cognitive one, for inspections of user-orientation, rather than of usability. A paper-based tool implementing this method has been developed and validated in the

[4] Part of this work has been carried out in the framework of the European Commission funded project eUSER ("Evidence-based support for the design and delivery of user-centred on-line public services", Contract no. 507180).

context of the eUSER project[5], and an online interactive version is also planned for later on. Future plans also include formal validation of the proposed framework throughout evaluation experiments of various types of systems, including mainstream systems developed following a Design for All approach, and systems dedicated to people with disability.

References

1. Savidis, A., & Stephanidis, C.: Unified User Interface Design: Designing Universally Accessible Interactions. In: International Journal of Interacting with Computers, 16 (2) (2004). 243-270.
2. eUSER: eUSER Conceptual and Analytical Framework (first version). In: Kevin Cullen (Ed), eUSER Deliverable D1.1, Part A (2004)
3. Rogers, E.: Diffusion of Innovations. New York: The Free Press (1993)
4. Fishbein, M., & Ajzen, I.: Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research, Addison-Wesley, Reading, MA (1975)
5. Benyon, D., Crerar, A., Wilkinson, S.: Individual Differences and Inclusive Design. In: Stephanidis, C. (Ed.) User Interfaces for All - Concepts, Methods, and Tools. Mahwah, NJ: Lawrence Erlbaum Associates (ISBN 0-8058-2967-9) (2001) 21-46
6. ISO 9241 ISO DIS 9241 – Part 11: Guidance on usability (1994)
7. Davis, F.: Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, 13, 3 (1989) 319-340
8. Featherman, M.S. & Pavlou, P.A.: Predicting e-Services Adoption: A Perceived Risk Facets Perspective. In: International Journal of Human-Computer Studies, 59, 4 (2003) 451-474
9. MORI: Measuring & Understanding Customer Satisfaction. A MORI Review for the Office of Public Services Reform. London: The Prime Minister's Office of Public Services Reform (2002)
10. Antona, M., Mourouzis, A., Kartakis, G., & Stephanidis, C.: User requirements and usage life-cycle for digital libraries. In: Proc. of the HCII2005 Conference (CD-ROM) (2005)
11. Mack, R. L., Nielsen, J.: Executive Summary. In Nielsen, J., and Mack, R.L. (Eds.), Usability Inspection Methods. John Wiley & Sons, New York, NY (1994) 1-23

[5] See `http://www.euser-eu.org/Document.asp?MenuID=123`

Basic Research of Input Support Device by Using Sympathetic Skin Response

Chikamune Wada and Fumihiko Masuda

Graduate School of Life Science and Systems Engineering,
Kyushu Institute of Technology
Hibikino 2-4, Wakamatsu-ku, Kitakyushu,
Fukuoka 808-0196, Japan
wada@life.kyutech.ac.jp,
masuda-fumihiko@edu.life.kyutech.ac.jp

Abstract. SSR is a biomedical signal, that reflects the activity of the sympathetic nervous system, and can be affected by cognitive thinking an decision making. We hypothesized that the SSR response could be used as a switch for an input support device for the ALS patients. In this paper, we investigated whether or not the SSR response could be used as a switch for an input support device for the ALS patients.

1 Introduction

Recently, many kinds of input support devices have been developed for the disabled. However, the residual motor function of a disabled person was necessary to operate most of them. If a disabled person has a progressive disease such as amyotrophic lateral sclerosis (ALS), it is necessary to modify the input support device according to residual motor functions. Whenever the input device is modified, the disabled person has to relearn how to use it. To overcome this problem, it is necessary to develop a device that would not depend on residual motor function.

One possible way to deal with this would be to use the sympathetic skin response (SSR) that can be obtained non-invasively and is independent from residual motor function. SSR is a biomedical signal, that reflects the activity of the sympathetic nervous system, and can be affected by cognitive thinking an decision making[1]. We hypothesized that the SSR response could be used as a switch for an input support device. However, there were still some uncertainties about using SSR. Specifically, there is no research whether or not respiration, visual/auditory stimulation, temperature or emotion affect the SSR. Among these factors, we have investigated the influence of respiration on the SSR [2].

Then, in this paper, first of all, we examined the influence of visual stimulation on the SSR. Next, removing the influence of respiration and visual stimulation, we investigated whether or not the SSR response could be used as a switch for an input support device for the ALS patients.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 429 – 436, 2006.

2 Influence of Visual Stimulation

2.1 Acquisition of SSR Signal

Fig. 1 shows the measurement system. Three Ag/AgCl electrodes were used to detect the SSR signal. The detection electrode was placed on the palm, the reference electrode was placed on the back of hand and the common electrode was on the wrist. The SSR signals were amplified 1000 times by an amplifier, and then inputted into a computer via an A/D converter. Fig. 2 shows a typical SSR signal pattern. We decided the latency and peak-to-peak duration as parameters of SSR signal.

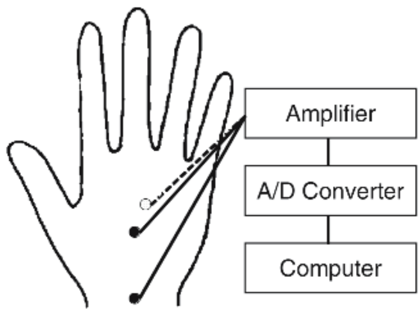


Fig. 1. Acquisition of SSR signal

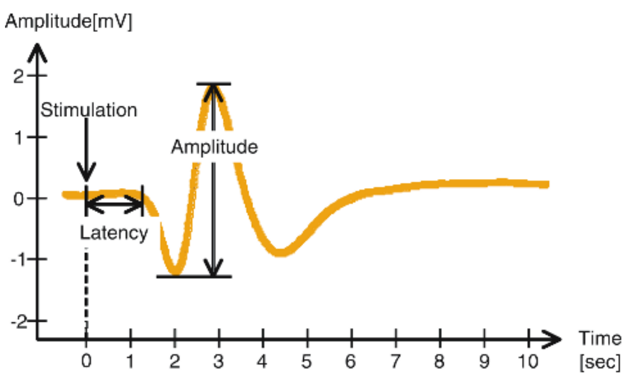


Fig. 2. A typical pattern of SSR and parameters for its evaluation

2.2 Experimental Method

Usually an input support device has some kind of visual display for the purpose of either inputting or getting information by the user. Since a problem arises when SSR responses involuntarily occur as a result of unintentional visual stimulus due to poor display coloring or lighting, we investigated the influence of visual stimulation on SSR occurrence.

Fig. 3 shows the experimental setup. Each subject was seated at a table facing a clock on it in a quiet room where the temperature was maintained between 25℃ and 28℃. An LCD screen was used to present visual stimuli to each subject. Subjects consisted of 4 able-bodied males (from 23 to 24 years old).

In this paper, three visual parameters were used. First was the color, second was the character size, and third was the character shape.

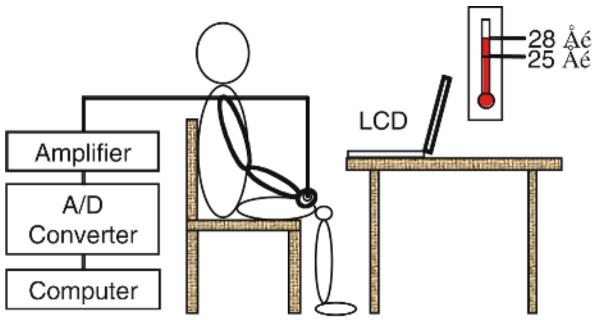


Fig. 3. Experimental setup

Color. The screen was adjusted to provide either a black image, or a color image. The colors chosen were blue, green, cyan, red, magenta, yellow and white. Testing was carried out using two luminance conditions. Under the first luminance condition the color brightness level was set to a predetermined level (the isoluminance condition) and under the second condition the brightness level was adjusted individually by each subject according to what they felt was the appropriate level (subjective isoluminance condition).

To avoid habituation, the order of the colors presented was random. Image interval time was 5seconds and this process was repeated 30 times for each color. During the trial subjects were only asked to view the image on the screen.

Size. In order to avoid the influence of character meaning, a rectangle was used instead of using character. The rectangles consisted of three sizes: 72*96, 288*384 and 384*512 pixels, while size of the screen was 1024*768 pixels. Color of rectangle was yellow and background color was black.

To avoid habituation, the order of the three rectangles presented was random. Image interval time was 5seconds and this process was repeated 20 times for each size. During the trial subjects were only asked to view the image on the screen.

Shape. In order to avoid the influence of character meaning, three kinds of geometrical figure was used instead of using character. Chosen figures were circle, triangle, diamond and square. Prior to participating in this experiment subjects were asked to choose the size of figure they felt was easiest for them to view. Then, the sizes were different in each subject. Color of figure was yellow and background color was black.

To avoid habituation, the order of the figures presented was random. Image interval time was 5seconds and this process was repeated 20 times for each color. During the trial subjects were only asked to view the image on the screen.

2.3 Results and Discussion

The SSR/visual stimulation appearance ratio was calculated by dividing the total number of times SSR signals were detected by the total number of times colors were presented. The relationship between SSR/visual stimulation appearance ratios and color is shown in fig. 4. Isoluminance condition results are indicated by square symbols, while subjective isoluminance condition results are indicated by triangular symbols. Fig. 4 shows the average for all results. SSR/visual stimulation appearance ratios for the subjective isoluminance condition was relatively lower than that for the isoluminance condition. It should be noted that there was little difference between SSR/visual stimulation appearance ratios amongst different colors under the subjective isoluminance condition. These results suggest that brightness levels had a significant influence on SSR response while differentiation between colors had little effect.

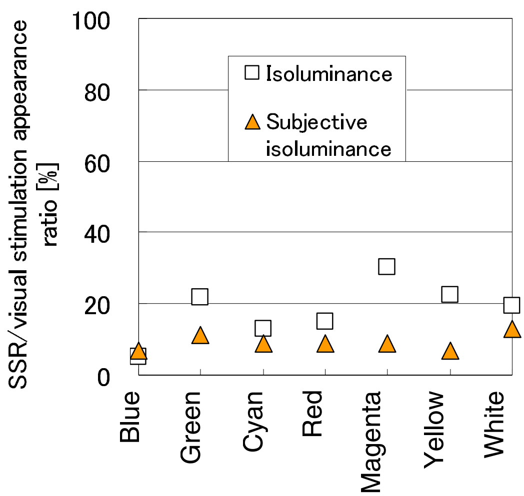


Fig. 4. Relation between SSR/visual stimulation appearance ratios and color

From this we concluded that the user would attain the optimal display conditions by first adjusting the display brightness for what he or she thought was the same level for all colors, and next, choosing the screen color that was easiest to look at.

The results of size and shape were shown in table 1 and 2. There were almost no different between sizes and shapes. Therefore, we thought that subjects should choose the character size that was easiest to look at and that there was no influence of character shape on the SSR.

Table 1. Relation between SSR/visual stimulation appearance ratio and figure size

Size (pixel×pixel)	72×96	288×384	384×512
Appearance ratio	2/80	1/80	3/80

3 Confirmation Experiment

In order to determine whether an input support device using SSR could be realized or not, we carried out an image viewing experiment to see if SSR could be effectively used as an input indicator for the ALS patient.

3.1 Experimental Condition

The experimental setup was the same as described in chapter 2. In this experiment, the subject was asked to imagine selecting a pre-determined target kana when it was displayed. Kana was one of Japanese characters. Forty-six different Japanese kana were individually displayed one after another for 3seconds each (see Fig. 5).

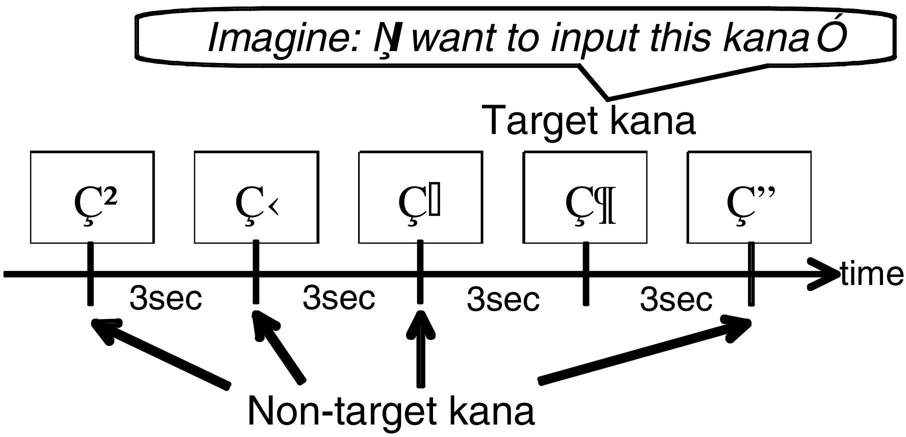


Fig. 5. Experimental procedure

For this experiment the subjective isoluminance condition was used and brightness levels for displayed kana was set according to the optimal display conditions arrived at in chapter 2. The color of kana was set to yellow with the background color being black. Prior to participating in this experiment subjects were asked to choose the color combination they felt was easiest for them to view. Most subjects selected the combination of a yellow kana on a black background.

As a result of the findings in past research [2], all of the SSR with latency under 1 second were excluded. Four non-disabled subjects and five ALS patients took part in this experiment. Non-disabled subjects consisted of three males and one female (from 22 to 24 years old). Profiles of the ALS subjects were shown in table 3. Experiment was continued until target kana was displayed 30 times.

Table 2. Profile of ALS patients

Sex	Age	Disease Duration	Condition
Male	65	16 yrs	Tracheotomy, A small quantity of finger movement
Male	67	14 yrs	Tracheotomy, A small quantity of eyebrow movement
Male	70	6 yrs	Tracheotomy, A small quantity of right limb movement
Male	63	11 yrs	Tracheotomy, A small quantity of left hand movement
Female	53	5 yrs	A small quantity of right leg movement

3.2 Results and Discussion

Fig. 6 and 7 show the amplitude and latency of SSR, respectively. In Fig. 6, the amplitudes of ALS subjects were significantly smaller than those of the non-disabled, except one ALS subject. It might be difficult to detect SSR signal when the amplitude was small. But, it would be possible to solve this problem because of improving measurement system. In fig. 7, the latencies of the non-disabled were significantly shorter than those of ALS subjects. However, it would not become a problem because the extended latency was 0.3msec.

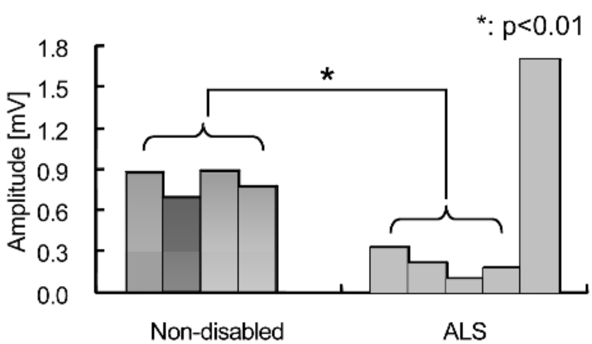


Fig. 6. Difference in amplitude of SSR between the non-disabled and the ALS subjects

The SSR/kana appearance ratio was calculated by dividing the total number of times SSR signals were detected by the total number of times kana displayed.

Fig. 8 shows the results of the SSR/kana appearance ratio. Bar graphs indicate the results when target kana was displayed, while diamond symbols indicate the results when non-target kana was displayed. Though the results varied amongst the different subjects, the ratios for the target kana were significantly higher than those for the non-target kana. Moreover, there was not significantly difference between the non-disabled and the ALS subjects.

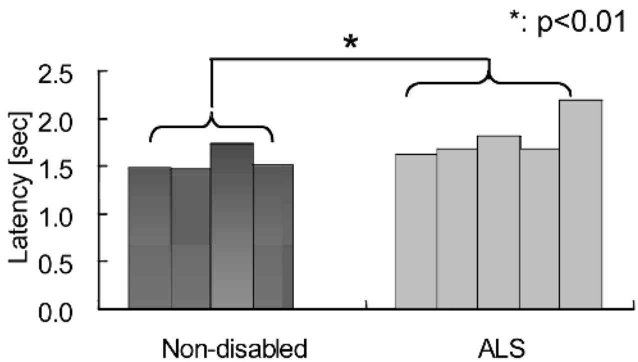


Fig. 7. Difference in latency of SSR between the non-disabled and the ALS subjects

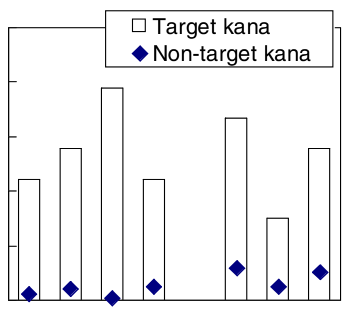


Fig. 8. SSR/kana appearance ratios for the target/non-target kana

It was previously reported that the cognitive process affects sympathetic nerve activity such as SSR (Miyagawa[3], Shimoda[4]). Therefore we believed that the reason that the SSR/kana appearance ratio increased was because sympathetic activity was activated when the subject perceived choosing the target kana. Then, we guess it is possible to develop an input support device using SSR as the switch for the ALS patients.

4 Conclusion

In this study, we investigated the influence of the visual stimulation on SSR in order to investigate whether it was possible to develop an input support device that used SSR as the device switch.

During the visual stimulation experiment, we found that there was little difference between the SSR appearance ratios amongst the various colors when subjects adjusted color brightness (luminance) to what they felt was the same level for all colors. Moreover, there was little difference between size and shape of character. As a result, we concluded that display luminance and character size should be adjusted subjectively.

Using these results, we carried out a character viewing experiment while measuring SSR. We found that the SSR appearance ratios were significantly higher when the subject was asked to perceive choosing a pre-determined target character compared with non-target character. Therefore, we conclude that this would allow for the possibility in developing an input device that could use SSR as the controlling switch.

Future plans are as below:

1. to determine the influence of environmental sound and emotion
2. to learn what role mental condition has on effectiveness

After solving these problems, we would like to develop an input support device for the disabled using SSR.

References

1. Iwase S., Mano T.: Sympathetic skin response (SSR), 1.Fundamentals, Clinical EEG electroencephalography, Vol.38, No.8, pp.558-565, 1996(in Japanese)
2. Masuda F., Wada C.: Basic Research into the Environmental Conditions Necessary to Develop a Sympathetic Skin Response Input Supporting Device, The Society of Instrument and Control Engineers Annual conference 2005, MP1-14-3(CD-ROM), 2005
3. Miyakawa M. and et al.: Sympathetic Skin Response and Event-Related Potential, The Autonomic Nervous System, Vol.29, No.5, pp.422-427, 1992(in Japanese)
4. Shimoda M. and et al.: Relationship between the P300 component of event-related potentials and sympathetic skin response in oddball and WS-IS paradigms, Japanese journal of electroencephalography and electromyography, Vol26, No.2, pp.120, 1998 (in Japanese)

Unskilled Finger Key Pressing and Brain Coherence

Ling-Fu Meng, Chiu-Ping Lu, and Ching-Horng Chen

Department of Occupational Therapy and Institute of Clinical Behavioral Science, Chang Gung University
259 Wen-Hua First Road, Gui-Shan, 333, Taoyuan, Taiwan
{lfmeng, cpl, godofwind72}@mail.cgu.edu.tw

Abstract. To press a computer key by an unskilled finger is sometimes an adaptive way to successfully access computer for the persons with quadriplegia. The efficiency of the unskilled site during the learning process should be addressed. Currently, we also want to know how the brain works in this unskilled situation during the learning process. Therefore, this combined motor behavioral and brain electrophysiological study was conducted. Since it was not easy to invite the persons with quadriplegia to participate electrophysiological studies, we invited eight typical college students to participate our study. Each of them tried to press the left, middle, and right keys for 200 times by their 2^{nd}, 3^{rd} and 4^{th} fingers respectively in a randomized order. The event-related coherence of the EEG was calculated to find out the functional connection among brain areas under unskilled (4th) and skilled (2nd) conditions. The result suggested that the alpha band synchronization between C3 and C4 electrodes under the unskilled condition was weaker than that under the skilled condition. It is likely that the performance of an unskilled finger was correlated to the weaker brain coherence. The brain might need some time to establish connections among different regions in the cortex during the learning process especially when using the unskilled control site.

1 Introduction

The persons with quadriplegia are constrained by their synergy motor pattern, desynchronized finger coordination, deformities and weak strength [1]. Some of them have to press computer keys by only one finger or some unfamiliar control site. Most often, the finger that they use are an unskilled and unusual one [2,3]. We want to know this finger's efficiency. Furthermore, knowing the relationship of the efficiency and the mechanism of brain is helpful to explain the characteristics of the performance for an unskilled finger while pressing a key [4,5,6]. Therefore, we combined motor behavioral and brain electrophysiological studies to substantiate the unskilled finger effect on the brain mechanism.

The relationship between coherence and motor efficiency was not clear and the results were often contradictory [4,5,6]. Deeny et al. addressed better motor efficiency of experts might be related to weaker brain coherence because the experts do not rely on their cortico-cortical connection to monitor motor task [4]. On the contrary, the findings of Knyazeva et al. indicated that the acallosal children had weaker brain coherence than the typical children under tapping conditions [5]. Being different from the explanation of Deeny et al., Knyazeva et al. found that poor weak brain coherence

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 437 – 441, 2006.

might be related to a poor motor performance. Furthermore, Serrien & Brown [4, 5, 6] substantiated that the brain coherence have changed dynamically with the motor behavioral performance while performing a bimanual task. Serrien & Brown emphasized that the learning stages should be seriously considered while trying to address the relationship of motor efficiency and brain mechanism.

At this moment, it is not convenient for the persons with quadriplegia to participate our brain electrophysiological studies. Therefore, some typical college students were recruited in our study. In order to select the matched control - skilled finger properly, it is also good to study typical persons since the 2nd finger can be treated as the controlled skilled condition comparing to the unskilled finger (4th finger).

2 Method

2.1 Participants

Eight right handed college students (2 males and 6 females) aging 19 to 24 (mean = 20.63, SD = 1.60) without any neuromuscular or cerebral disease participated in this study. The averaged handedness quotient of self reported Edinburgh handedness inventory was 93.75 (±10.61).

2.2 Variables

We took the individual finger (index, middle, or 4th finger) which is used to press the computer key as an independent variable. Dependent variables included the reaction time (the period between seeing the number and the action of pressing the corresponded key) and the brain coherence (Fig 1) between C3 and C4 electrodes〔Coh (C3, C4)〕in the 8-10 Hz alpha band. The value of coherence was between 0 and 1.

2.3 Experimental Design

The experimental research design was used to answer the research questions. Each participant used right hand's fingers to press the computer keys. If the number of 2 was seen on the screen, the participants used their index finger to press the corresponding key on the left side of the keyboard. Respectively, they pressed the corresponding key in the center of the keyboard with the middle finger and the key on the right side with the 4th finger if 3 or 4 was seen on the screen. Each finger pressed the keys for 200 times and the order of the 600 attempts was totally randomized (Fig 1).

2.4 Stimulus Presentation and Key Pressing Performance

The timing of the stimulus presentation was controlled and subject responses (accuracy and reaction time) were recorded using Stim II Software (Neuroscan, Inc. Sterling, VA, USA). The stimuli included Arabic number 2, 3, and 4.

2.5 Electroencephalogram (EEG) Acquisition and ERP Recording

EEGs were recorded from Quick-Cap 32 Channel-Sintered electrodes located in the standard 10-20 system. All electrodes impedance were brought to below 10 kΩ. The

EEG was band pass filtered (1-30 Hz) and digitized at a sampling rate of 1000 samples/s. The baseline for ERP measurements was the mean voltage of a 100ms prestimulus interval. Attempts exceeding±100μV at horizontal and vertical electrooculogram (EOG) were excluded immediately. Furthermore, attempts with eye blinks, eye movement deflections, and over ±60μV at any electrode were also excluded from ERP averages.

2.6 Procedure

Each participant used right hand's fingers to press the computer keys. If the number of 2 was seen on the screen, the participants used their index finger to press the corresponding key. They pressed the keys using their middle or 4th finger if they saw the number 3 or 4 respectively. Each finger pressed the key for 200 times (attempts). The order of 600 attempts were totally randomized (Fig 2)

2.7 Statistics

The multiple Wilcoxon's z tests were conducted to find out the unskilled finger effect on reaction time and the coherence between C3 and C4.

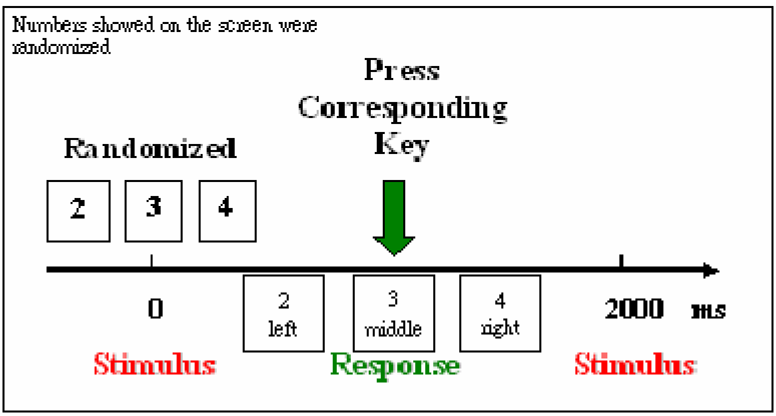


Fig. 1. The participants used their index, middle, 4th fingers to press the corresponding keys respectively when seeing number 2, 3, 4. The order of 600 attempts (200 attempts for each number) was totally randomized.

3 Results

The data of the results of the 8 participants were listed in the Table 1. The Coh (C3, C4) under the skilled condition was stronger than the unskilled condition (Z = -0.840, *p* = .045, one-tailed). However, three participants demonstrated the situation that the action of the 2nd finger was slower than the 4th finger. Therefore, for those three, the 4th finger could not be treated as the unskilled condition. In order to veritably compare the skilled and unskilled condition, we excluded those 3 participants' data and further analyzed the Coh (C3, C4) of the remaining 5 participants whose 2nd finger performed exactly faster than the 4th finger. Table 2 showed that the reaction time was significantly slower and the Coh (C3, C4) was weaker under the unskilled condition comparing to the skilled condition (Table 2).

Table 1. A comparison between 2nd and 4th finger (8 participants)

	Conditions (Mean ± SD)			
	2nd finger (skilled condition)	4th finger (unskilled condition)	*Wilcoxon's* Z	*P* (one-tailed)
Reaction Time	463.03 ± 37.48	474.39 ± 35.18	-1.690	.205
Coh (C3,C4)	0.432 ± .232	0.317 ± .133	-0.840	.045

Note. Coh (C3, C4) means the coherence strength between C3 and C4 electrode.

Table 2. Comparison between 2nd and 4th finger (after excluding 3 participants whose 4th finger performed faster than the 2nd finger)

	Conditions (Mean ± SD)			
	2nd finger (skilled condition)	4th finger (unskilled condition)	*Wilcoxon's* Z	*P* (one-tailed)
Reaction Time	465.570 ± 46.531	489.122 ± 33.665	-2.023	.022
Coh (C3,C4)	0.486 ± .284	0.317 ± .160	-2.023	.022

4 Discussion

In this study, the alpha band coherence value of the 2nd finger was stronger than the 4th finger. Therefore, similar to the finding of Knyazeva et al.[5], the relationship between brain coherence and pressing efficiency did exist after comparing the coherences under skilled and unskilled condition. After analyzing the 5 participants whose 2nd finger performed really better than the 4th finger, the coherence became stronger under the skilled condition (pressing with 2nd finger). Therefore, we further confirmed the argument made by Knyazeva et al.[5] that the unskilled and skilled finger might be related to weaker and stronger brain coherence respectively.

We also agree with the dynamic view proposed by Serrien et al.[6]. During the course of pressing, we found the brain coherence became stronger and weaker in the fatigue stage. Although the coherence was weaker under the unskilled condition intially, the stronger coherence would possibly occur after the efficiency was improved.

Although the finding in this study was different from the result of Deeny et al.[4], the argument proposed by Deeny et al. was also reasonable. The relation between weaker coherence and better performance in the experts mentioned by Deeny et al. should be substantiated under the situation of a very easy or familiar status. In our study, the motor response to the task demands had not been considered as an easy or familiar way for both 2nd finger (skilled) and 4th finger (unskilled conditions) since the keys were not easy to be pressed. Therefore even the skilled 2nd finger could not be treated as an expert situation mentioned in the study of Deeny et al.[4]. To conclude, if the two conditions (ex. 2nd and 4th finger) were both in an unfamiliar context, a better condition (2nd finger) could exhibit a stronger brain coherence to reflect that a better condition still occupies an advantaged position during the learning process.

References

1. Ling Fu Meng, Tien-Yu Li, Chi Nung Chu, Ming-Chung Chen, Chien-Huey Sophie Chang, et al.: Applications of Computer Access Approach to Persons with Quadriplegics. Lecture Notes in computer Science, 3118, 857-864 (2004)
2. Turpin G, Armstrong J, Frost P, Fine B, Ward C, Pinnington L.: Evaluation of alternative computer input devices used by people with disabilities. Journal of Medical Engineering Technology, 29(3):119-29 (2005)
3. Ding D, Cooper RA, Kaminski BA, Kanaly JR, Allegretti A, Chaves E, Hubbard S.: Integrated control and related technology of assistive devices. Assistive Technology, 15(2):89-97 (2003)
4. Deeny, SP; Hillman, CH; Janelle, CM; Hatfield, BD.: Cortico-cortical communication and superior performance in skilled marksmen: An EEG coherence analysis. Journal of Sport & Exercise Psychology, 25(2), 188-204 (2003)
5. Knyazeva, M; Koeda, T; Nijokiktjien, C; Jonkman, EJ; Kurganskaya, M; de Sonneville, L; Vildavsky, V.: EEG coherence changes during finger tapping in acallosal and normal children: A study of inter- and intrahemispheric connectivity. Behavioural Brain Research, 89(1-2), 243-258(1997)
6. Serrien, Deborah J; Brown, Peter.: Changes in functional coupling patterns during bimanual task performance. Neuroreport: For Rapid Communication of Neuroscience Research, 15(9), 1387-1390 (2004)

Customizing User Interfaces with Input Profiles

Mario Mühlehner

Institute Integriert Studieren, University of Linz,
Altenbergerstrasse 69, A-4040 Linz, Austria
mario.muehlehner@jku.at

1 Introduction

In this paper I will present an approach to design custom user interfaces in order to tailor the human-machine-interaction to individual user skills and preferences. These so called *input profiles* play a crucial role in the *HeadControl+ framework* [2].

HeadControl+ (which is now in its integration stage before extensive system and user testing) represents our efforts in building a new type of user interface that is highly affordable and yet extremely customizable to the needs of its users. The user wears a laser unit on its head (by means of a headset or clipped to his/her glasses), which laser beam points to the center of a navigation area located near the computer display (typically in front of the display where normally the keyboard resides). A non-expensive web cam mounted above the navigation area (which roughly has the size of the computer display) observes the laser movements, thus enabling the HeadControl+ software to extract the laser point coordinates. Whenever the user turns his/her head, the laser point performs a corresponding, horizontal motion. Bowing and raising ones head leads to vertical movements of the laser – thus the system works like a typical head tracker. After the stream of laser coordinates is extracted utilizing image analyzing algorithms, it is used to control the computer. Therefore *virtual devices* have been conceived; a virtual device is a software component that processes the laser interaction (and optionally the operation of connected *input sensors* like simple switches, pneumatic switches, etc.) in order to emulate a conventional input device (e.g. mouse) or even perform complex actions (play back of text macros, starting programs, etc.).

Each virtual device can be perceived as a separate, exclusive input mode. Whereas the original HeadControl system [1] was only able to control the mouse cursor via head movements, HeadControl+ represents a *multi modal* input interface. This multi modal approach urges for further research (for instance: which meaning has the activation of a switch in a particular virtual device?) that shall be discussed subsequently.

2 State of the Art in Research

In general, multi modal user interfaces comprise distinct communication channels between the user and the computer system. The most common means of communication are (besides standard input devices like mouse and keyboard) speech, gestures, eye or head tracking, etc. [4,5]. The related research has its focus mainly on how to enhance user interfaces (in particular for assistive technology) by combining two or more

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 442 – 449, 2006.

communication channels in a system-theoretical approach. Despite there are efforts to integrate distinct controls into one common user interface, for instance the joystick on an electrical wheel chair as generic control to manipulate the environment (TV, lights, etc.) [6], it remains unclear *how* to model the relations between the usage of the controls and its executed actions in the different contexts respectively modes. Wang and Mankoff [7] address this issue from an information-theoretical point of view and show how to build bridges between different types of input devices (for instance to emulate a conventional mouse via mini keyboard). They point out that simple adaptor applications can be designed using scripts, more sophisticated applications need to implement software components. The presented approach focuses only on translating the occurring events of one input device (e.g. a key stroke on a keyboard) into events of a target device (e.g. mouse movements). Because the input devices of HeadControl+ are literally *virtual*, a restriction to conventional device events or functions seems not reasonable.

3 Input Profiles

In the master thesis *“HeadControl Plus – A multi modal input device for physically impaired computer users”* [3] a novel approach has been presented. The basis of the concept is the insight that conventional input devices have a strict relation of cause and effect (or in other words *input event* and *executed action*). For instance, the left mouse button is supposed to cause an according “left mouse button pressed” event whenever it is pressed. There is no reasonable setting (or context) when pressing the left mouse button should provoke another effect.

When speaking of multi modal user interfaces which make use of the same controls (e.g. input sensors) in different contexts, it is neither clear what pressing a switch in one context will effect, nor in any other context (or mode). And because interface hardware (special hardware that allows to connect input sensors to a computer) and input sensors are quite expensive, it is desirable to reduce the number of input sensors needed to operate a user interface. Another reason to keep the number of required prerequisites low is that a considerable percentage of physically impaired users might not be able to operate more than one or two input sensors.

3.1 Basic Concept

The basic concept distinguishes between *event sources* and *event sinks*. An event source fires events each time an event of a particular type occurs (for instance when an input switch is pressed). Thus event sources generate a stream of (input) events which can be directed to event sinks in order to be processed in an appropriate way.

The simplest event type is a notification: it holds no additional information except the implicit message “something has happened”. Imagine a binary input switch capable of being pressed and released (like a key on the keyboard). The input switch can be perceived as a trigger containing two event sources: one representing the *“pressed”* event (called *Button down*) and one representing the *“released”* event (called *Button up*). Each time when the input sensor is pressed, an event is generated by its software equivalent and fired by the *Button down* event source; when it is released, an event is fired by the *Button up* event source. The software component modeling the input switch is called *trigger unit*, because it comprises only event sources, thus being able to trigger arbitrary actions.

An entity that aggregates only event sinks (respectively actions) is called *execution unit*. All functions associated with a conventional mouse (move, press/release buttons) for instance can be modeled in an equivalent execution unit featuring an event sink for each function.

Dispatcher units represent besides the trigger and execution units the third type of *input profile design components* – they contain event sources as well as event sinks and are conceived as transformation units which process events (on its event sinks) and generate new events and fire them (on their event sources). A dispatcher unit has one or more event sinks and minimum one event source.

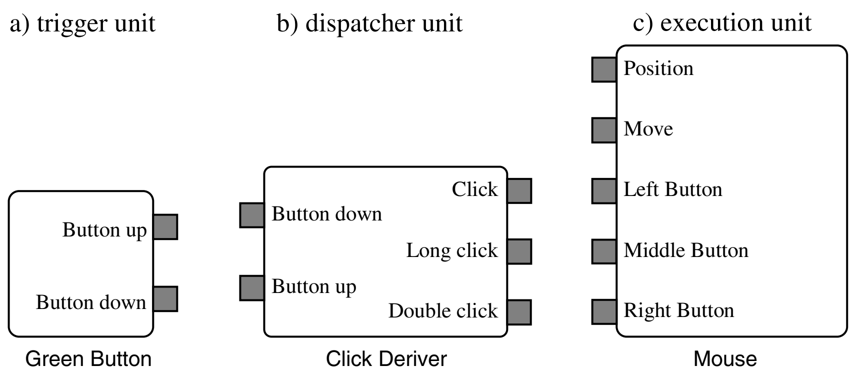


Fig. 1. Types of design components

Fig.1 shows examples for the described design components; each design component is rendered as a black box with one or more connection points representing the event sources/sinks. Component 1.a models a binary input switch. Component 1.b shows a dispatcher unit intended to be connected to an input switch trigger unit. It measures the time between pressing and releasing the connected switch and classifies the input (related to the double click time defined by the OS) into a simple click, a double click or a long click (the switch remains pressed during observation period). Thus the *Click Deriver* contributes to a refined input event handling. Component 1.c is an example for an execution unit: the *Mouse* component, supporting basic mouse functions.

Fig. 2 pictures the two fundamental ways to connect event sources and event sinks. Fig. 2.a represents the case of connecting one event source to exactly one event sink, which means that the events fired by the event sink are provided to the connected event sink exclusively. Fig. 2.b (as implemented in HeadControl+) allows connecting more than one event source to an event sink, therefore implementing a logical *OR*: the action represented by the event sink is executed if event source 1 *or* event source 2 … *or* event source *n* fires an event. Thus redundant accesses to an action can be defined avoiding ambiguities (as it would be the case when one event source would be allowed to connect to more than one event sink: in which order shall the fired events be handled by the connected event sinks?).

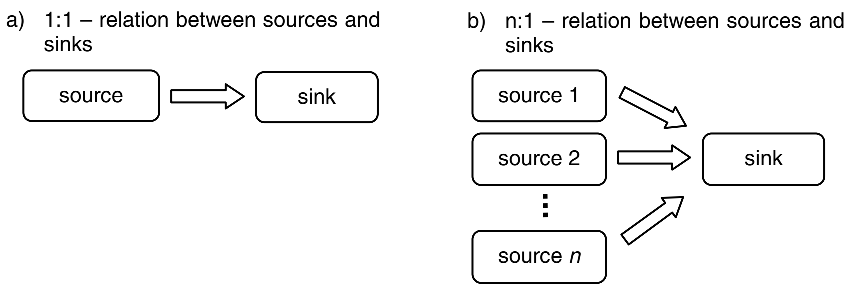


Fig. 2. Basic connection types between event sources/sinks

Connecting event sources to event sinks finally leads to a directed graph; the nodes are represented by design components (trigger units, dispatcher units, execution units), while the edges denote the flow of events within the graph. The sum of all design components and its connections needed to describe a specific interaction scenario is called *input profile*.

3.2 Framework Support

HeadControl+ has been conceived as a plug-in framework with hotspots to add new virtual devices, new input sensor types, new design components (hotspots for trigger units, dispatcher units and execution units) as well as new event objects to be communicated between event sources and event sinks (there are a few more hotspots not mentioned here as they are not relevant for the subsequent discussion).

The plug-in architecture has been chosen for various reasons; the most important one is the immense openness and scalability it provides. Since the design components are conceived as building stones to model the behavior of individual interaction scenarios, the extensibility of HeadControl+ is crucial to effective interface modeling.

Each design component manages its own collection of properties which can be customized by the user interactively; it contains at least the label of the component. Finally, design components can be enabled and disabled. When disabled, a design component does not respond to incoming events nor is it sending any events. Each design component is responsibly for persisting its settings.

Virtual devices are implemented as plug-ins, too. Like any design component a virtual device can be enabled and disabled (for instance when switching from one virtual device to another). The core of a virtual device is an input profile; it specifies which resources of the input interface are involved and how they are connected to individual input actions.

Finally, the *input profile editor* is a graphical editor and also part of the HeadControl+ framework. It has been conceived to design input profiles interactively – just by inserting design components from the component library into a new input profile and connecting the event sources and sinks via dragging from one connection point to another.

3.3 Design Components

HeadControl+ comprises already a fundamental set of design components providing the design of quite complex interaction schemes. To the set of trigger units belong the *LaserMotion* unit – which provides the movements of the laser point within the navigation area as well as events when the laser point disappears or enters the navigation area again – and the *VirtualToolBar*, a customizable toolbar containing an arbitrary number of buttons and button-shaped checkboxes which manipulation fires again events (that can be used for instance to build a menu-based mouse click assistant to perform double clicks, drag & drop, etc.).

The actions provided by the currently available execution units fall into three categories:

a) functions associated with conventional input devices (mouse, keyboard)
b) HeadControl+ control commands
c) miscellaneous functions

The execution unit *Mouse* offers functions to relocate the mouse cursor (both relative and absolute) as well as the opportunity to access the mouse buttons, while the *Keyboard* execution unit receives events containing the key codes of the keystrokes to be generated. The *DeviceSwitcher, SignalGenerator* and *LevelController* execution units are intended to interact directly with the HeadControl+ system and the interface hardware it manages. The *DeviceSwitcher* allows to activate other virtual devices (for instance in order to switch from the mouse emulation to the keyboard emulation), while *SignalGenerator* and *LevelController* are conceived to render signals on the interface hardware (for instance flashing LEDs or adjusting the intensity of the laser beam).

The third category of execution units consists of components to enhance the user convenience; it comprises the *KeyboardMacro* execution unit (which is able to execute complete pre-recorded keystroke or text sequences), the *ApplicationLauncher* execution unit (which is used to start programs like the Windows™ Explorer or any other desired executable) and the *AudioPlayer* (which plays a selected audio file whenever it receives a notification event).

The set of dispatcher units consists of generic components intended to fit in as much scenarios as possible. The *ModuloCounter* for instance is a counter that increases its value whenever a notification reaches its *Increase* event source. The amount of increment can be customized by the user. In addition, the *ModuloCounter* allows the specification of a maximum count. Whenever the maximum count is exceeded, an appropriate event is fired on the *Overrun* event source (another event source fires notifications each time a counter underrun occurs). Setting the maximum count of a *ModuloCounter* to the value of two (and the increment to the value one) can be used to accomplish a toggle unit that switches between the states *0* and *1* (depending on the processed *Increase* events).

Multiplexer dispatcher units serve as routers for incoming events; they have one input to receive events (event sink *Input*) and multiple outputs representing possible destinations for these events (event sources *Output* 1..n). By sending an event holding the number of the desired destination output to the *Sel* event sink of the *Multiplexer*, all following events received at the *Input* event sink are routed to the specified destination event source. The *DeMultiplexer* performs the inverse operation of the *Multiplexer* – it collects events on its various inputs (event sinks *Input* 1..n) but dispatches only the ones at the specified event sink to its event source (event source *Output*).

These are some examples of design components already provided by the HeadControl+ framework, among the ones not mentioned are components to explore additional interaction sources (*GestureDetector*), components to offer alternative pointing metaphors (*JoystickConverter*), etc. Many more are expected to follow.

3.4 Examples

Fig. 3 shows a simple input profile to emulate a two-button mouse. There are two input switches involved (named *Green* and *Orange*). The event source *Up* of input switch *Orange* is directly connected to the *Right Button* event sink of the *Mouse* execution unit, thus executing a single click of the right mouse button each time the *Orange* switch is released (the button event sinks of the *Mouse* execution unit are able to process two kinds of events: when receiving a notification without any parameters they execute a single click of the corresponding button by default; when receiving an event with one numeric parameter, a “button down” action is performed on value *1* and a “button up” on value *0*). Input switch *Green* is connected to the *Left Button* event sink of the *Mouse* execution unit including two *Weighter* dispatcher units. The *Weighter* components convert the incoming notifications to events with numeric parameters (*0* respectively *1*), thus sending a “0”-event to the *Left Button* event sink when *Green* is pressed, a “1”-event when it is released. This way the left mouse button is fully associated to the *Green* input switch and can be used to select multiple objects by holding the *Green* input switch while moving the mouse pointer.

The *LaserMotion* component provides the normalized coordinates of the laser point. By directing its *Move* events (each holding the coordinates of the current laser point position) to the *Position* event sink of the *Mouse* execution unit an absolute mapping of the laser movements to mouse pointer movements is accomplished.

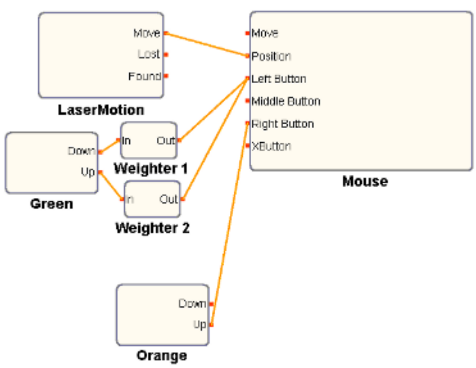


Fig. 3. Input profile emulating a two-button mouse

Fig. 4 shows a more complex input profile emulating a virtual keyboard that is capable of relocating the mouse cursor without switching to another virtual device. During regular operation the user selects virtual keys painted at his/her navigation area and confirms his/her choice by pressing the *Green* input switch. Doing so, the user is able to enter text into a textbox (which is for instance part of a form). Normally a user would – in order to focus another textbox – switch to the virtual mouse, move the mouse cursor over the desired control, click it (to set the input focus on it) and switch back to the virtual keyboard to continue text input. This input profile implements two *sub-modes* that are actually triggered by the *Orange* input switch. As long as the user holds the *Orange* input switch, the profile remains in the “cursor relocation mode” enabling he/she to control the mouse cursor with his/her head movements. When releasing the *Orange* input switch, the mouse cursor freezes at its previous position and an automatic left button mouse click is generated subsequently (setting the input focus to the control beneath the mouse cursor). Afterwards the input profile is in “keyboard mode” again allowing the user to continue text input.

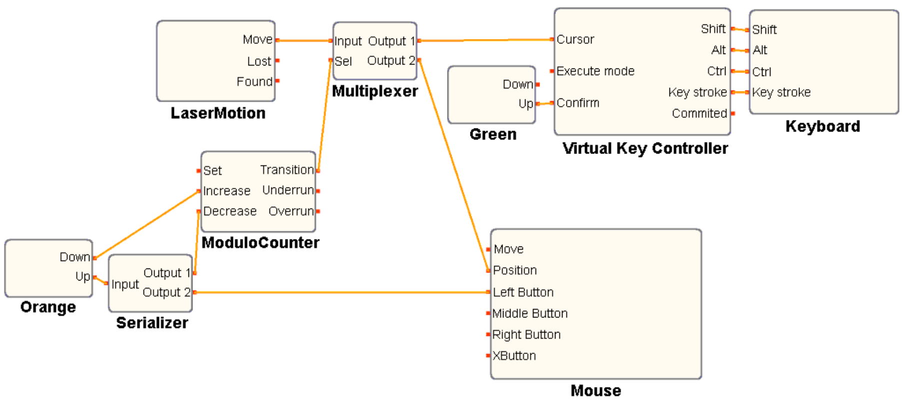


Fig. 4. Input profile emulating a keyboard (plus mouse)

This behavior is accomplished by employing the *ModuloCounter* to toggle the current mode of the input profile. Depending on its value (*0* or *1*) the *Multiplexer* redirects the incoming laser point coordinates (provided by the *LaserMotion* unit again) to the *Virtual Key Controller* (regular keyboard mode) or to the *Mouse* execution unit (cursor relocation mode). The *Virtual Key Controller* is the central component of the input profile; it holds the layout of the current virtual keyboard and controls how virtual keys are accessed (in this case the *Green* input switch connected to the *Confirm* event sink determines the confirm mode). Keyboard strokes associated to virtual keys are delegated to the *Keyboard* execution unit (as well as the states of the modifier keys *Shift*, *Alt*, *Ctrl* which are also part of the virtual keyboard layout and managed by the *Virtual Key Controller*).

The *Serializer* dispatcher unit is connected to the *Up* event source of the *Orange* input switch. It duplicates the incoming notification events each time the input switch is released and broadcasts them on its outputs in ascending order with a customizable time delay, thus setting the *ModuloCounter* value to *0* first and performing a left button mouse click afterwards.

4 Conclusions

HeadControl+ is a novel non-classical user interface that takes advantage of multi modal interaction in at least three ways. Firstly, it provides multiple virtual input devices, each representing an exclusive interaction mode. Secondly, the input profiles - which build the core of each virtual device – make use of generic design components introducing distinct sources of user input: the pursuit of laser interaction in order to track the user's head movements, the operation of various input sensors (switches, etc.), derivation of pointing gestures, etc. Since the plug-in architecture of HeadControl+ supports the continuous extensibility of the framework, future development will consider the application of speech recognition to enhance user interfaces further.

Last but not least, input profiles allow modeling sub-modes to design even more complex interaction schemes reducing the need to switch between virtual devices repeatedly, thus increasing convenience during operation. Input profiles can be easily created. Based on a comprehensive library of design components, the input profile editor application allows designing input profiles without any coding.

In conclusion, input profiles can be tailored quickly to the individual skills and preferences of computer users, thus turning out to be a versatile and effective approach to address the needs of physically impaired computer users.

Acknowledgement

This work is supported by the Fund of the Oesterreichische Nationalbank.

References

1. Mühlehner, M.: „HeadControl – Ein lasergesteuertes Eingabegerät für körperlich beeinträchtigte Computeranwender“, Diploma Thesis, Universität Linz, 2003.
2. Mühlehner, M., Miesenberger K. „HeadControl+: A multi modal input device“, Pages 204-211 Computers Helping People with Special Needs Conference 2002, ed. Klaus Miesenberger, Joachim Klaus, Wolfgang Zagler, Springer Verlag 2004.
3. Mühlehner, M. “HeadControl Plus – A multi modal input device for physically impaired computer users”, Master Thesis, Universität Linz, 2005.
4. Horvitz E., Shwe M. “In Pursuit of Effective Handsfree Decision Support: Coupling Bayesian Inference, Speech Understanding, and User Models”
5. Oviatt S. “Multimodal interfaces”, Handbook of Human-Computer Interaction, ed. J. Jacko, A. Sears, Lawrence Erlbaum, New Jersey, 2002.
6. Van de Loo S., Symonds B. “One Control Everywhere for Everything”, Australian Rehabilitation & Assistive Technology Association Conference, 2004.
7. Wang J., Mankoff J. “Theoretical and Architectural Support for Input Device Adaption”.

Computerized Assessment Approach for Evaluating Computer Interaction Performance

Ming-Chung Chen[1], Chi-Nung Chu[2], Ting-Fang Wu[3], and Chih-Ching Yeh[2]

[1] National Chiayi University, Department of Special Education,
300 University Rd., Chiayi, Taiwan
mtchen@mail.ncyu.edu.tw
[2] National Taiwan Normal University, Department of Information & Computer Education
162, Sec. 1 Hoping East Rd., Taipei, Taiwan
nung@mail.ckitc.edu.tw,
dexteryeh@mail.cy.edu.tw
[3] National Taiwan Normal University, Graduate Institute of Rehabilitation Counsel
162, Sec. 1 Hoping East Rd., Taipei, Taiwan
tfwu@ntnu.edu.tw

Abstract. This study presents a computerized assessment approach for evaluating a subject's pointing and selecting proficiency using computer input tools, to aid access tool selection for users with severe disabilities. The CAT system consists of three subsystems. The CAT system not only provides clinicians with an objective means of evaluating clients' specific mouse operating difficulties, but also allows them to compare the performance improvement made by a client make during the device selection and training period. The client's performance in each assessment task is assessed on the basis of speed, accuracy and efficiency. Besides introducing the CAT system, this study also describes an example of adopting the CAT system to assist a client to select a suitable pointing device.

1 Introduction

Computer and Information Technology (CIT) plays an increasingly important role in daily life. CIT can assist physically challenged individuals to participate in educational, vocational, societal and other daily activities in many ways. Moreover, CIT is considered as an "equalizer" for people with disabilities to participate fully in e-society. However, older people or people with disabilities encounter various difficulties when interacting with computers, and need extra adaptive solutions to help them to interact with CIT.

Many adaptive strategies for computer access and various alternative input devices are available for older people and people with disabilities [1]. Selecting suitable strategies and devices for a client is difficult for rehabilitation professionals, partly due to the lack of an adequate assessment tool to evaluate a client's performance of interacting with computer. Clinicians need related methods to assist them to execute a proper computer interaction evaluation. Clinicians require tools to assess the client's mouse operating capability during the evaluation process and determine suitable access methods; to compare the performance of feasible devices, and to understand the change of client's performance during training program delivery.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 450 – 456, 2006.

Some assessment approaches, including checklist, flow chart, evaluation tasks battery and software, are available for clinical use[2,8,9,10,11]. Checklists and flow charts focus on an assessment procedure that provides systemized evaluation steps and corresponding strategies or devices. An evaluation task battery provides a group of essential mouse operating tasks. Evaluators score a client's mouse operating performance by asking clients to complete some tasks and judge by themselves. Evaluation software allows clients to complete mouse operating tasks generated by software and collect clients' response automatically. Many such software systems have been developed, but mostly for specific research purposes. Moreover, these computerized assessment tools assess pure mouse proficiency instead of functional mouse operating performance. Evaluating clients' functional mouse operating performance provides practical information for both clinical rehabilitation professionals and clients. This information is very important for selecting the most appropriate device for interacting with computers in authentic situations.

In addition to the essential evaluation tasks, previous studies also considered what information data should be collected and analyzed is also an important consideration in the past researches. Speed, accuracy and efficiency are regarded as essential indicators of mouse operating performance. Recent studies have particularly emphasized efficiency as a fundamental indicator. Efficiency is defined in terms of stability when moving the cursor, such as the inconsistency between the trajectory of cursor movement and ideal straight line, movement variability, and sub-movements appeared during movements [2,5,6]. The qualitative data helps clinicians understand the effect of different factors, such as distance, direction and device, on cursor movements.

Our previous study had previously presented a Computerized Assessment Tool for Mouse Proficiency (CAT-MP) [4]. The authors aim to develop an integrated evaluation system based on the features of CAT-MP, containing pure mouse proficiencies and functional interaction skills, as well as measuring important indicators mentioned above.

2 CAT System

To develop a system containing essential mouse proficiencies and important functional interactions, this study reviewed related literature and analyzed the tasks or components of common interactions in advance. A checklist for conducting an investigation was drafted from the analytical results. The checklist has two parts, one focused on mouse proficiencies which contained static clicking, cursor moving, pointing and clicking, and dragging; the other one focused on functional interactions that comprise Operation System manipulating, window operation and word processing.

Eleven professionals from special education, rehabilitation, computer education, and computer science participated in the investigation, by ranking the degree of importance of each task in measuring the above proficiencies and interactions. The responses indicate that the professionals regarded most of the tasks as "important". These important tasks were included in the Computerized Assessment Tool (CAT), which contains three sub-assessment systems, namely Basic Skills (CAT-BS), Mouse Proficiency (CAT-MP) and Functional Performance (CAT-FP).

Based on the design idea, CAT-BS focuses on essential mouse proficiency testing, and provides standard evaluation tasks and procedures. The major aim of CAT-BS is to evaluate the fundamental mouse proficiencies. Evaluation results can be adopted to compare a user's performance with the average performance of all users. CAT-MP enables clinicians to set up various testing situations to explore the specific difficulties and suitable environments when client interacts with computer, such as appropriate interval between double-clicking, proper size of target for clicking, preferred color of the background. The CAT-FP system is used to determine the performance of multi-step functional interactions after equipping the client with some proper devices.

2.1 System Overview

CAT consists of three modules, namely client's Basic Data Module, Assessment Module and Analysis Module. The Basic Data Module records clients' related data, including computer usage needs, capability of sensory, motor control, and cognition, device used for assessment. The Assessment Module comprises assessment program and corresponding database. The Assessment program allows clinical evaluators to set up assessment tasks, and tests the client performance. The database records the client's detail responses, and gives data to the Analysis Module to understand the client's performance and difficulties.

2.2 Assessment Module

As mentioned earlier, the Assessment Module has three sub-assessment systems in. Each sub-assessment system has its own specific purpose and assessment items.

CAT-BS. CAT-BS system only comprises three tests, each with one task to evaluate. CAT-BS involves "targeting and left-single clicking", "targeting and left-double clicking" and "dragging". Each test comprises 32 tasks generated from four distances (1 cm, 5cm, 10cm, 15cm), each with eight directions (0, 45, 90, 135, 180, 225, 270, 315). The "targeting and left-single clicking" test requires clients to move the cursor from an initial icon to the target icon and activate left-single clicking. The "targeting and left-double clicking" test assesses the cursor moving and left-double clicking performance. The "dragging" test asks the user to move an icon to a target area by maintaining the left click down on a target icon, and releasing it on when moving it in the target area.

Although the program does not permit parameter setting, and automatically controls the assessment procedure, CAT-BS can be utilized to examine clients' basic mouse operating skill and compare the performance of various interaction devices or preintervention and prointervention.

CAT-MP. The CAT-MP system encompasses four tests, each consisting of three evaluation tasks, namely "targeting", "stationary clicking", "targeting and clicking" and "dragging". Within each test, the evaluator can flexibly set up the target size, target color, targeting distance, targeting direction, amount of testing of each task and time permitted for each trial. Test tasks can be arranged uniquely according to the client's proficiency or aim of the assessment. For instance, a dragging task of 15cm and 5cm moving distance with a 1cm^2 icon is chosen if the assessment measures the distance that a client can drag. Conversely, the evaluator could fix the moving distance and change the icon size to explore the most suitable size for a client to click the mouse button.

In the “targeting” test, the client has to move the cursor to the specified icon and keep the cursor in the icon for a second. This test explores the performance of moving cursor in different distance and directions.

“stationary clicking” measures the subject’s stationary clicking performance regardless of clicking device. Without moving cursor, a subject must click the target icon by pressing the button. The test results not only investigate the button activation situation, but also detect the proper interval of time between two clicks. This approach is useful for setting the double-click time interval.

“targeting and clicking” studies the ability to move the cursor and click coordinately. The “dragging” test aims to measure the client’s dragging performance. The client is asked to drag the target icon by maintaining the button down, and to drop it when the target icon is located in the destined area. These two tests are performed when the client could move cursor and clicks the button, and examine a user’s coordination when moving the cursor and pressing a button either sequentially or simultaneously. For instance, a possible test is to move the mouse cursor to a destined area 15cm away in eight radial directions with a trackball, and activate a single switch to click inside a 1 cm^2 icon, to determine whether he could move the cursor to the destination and simultaneously click the target icon by a trackball with a left-click switch.

CAT-FP. Unlike the aforementioned two systems, CAT-FP focuses on typical interactions in a real Windows system. CAT-FP comprises three tests, namely “Operation System manipulating”, “window operating” and “word processing”. Each test involves several typical evaluation tasks. The “Operation System manipulating” test includes manipulating the "start menu", copying an icon and pasting it, executing a file on desk and shutting down system. Every test task requires many subtasks to be completed in multiple steps. For example, when a client wants to complete the task of “manipulating "start menu"”, he needs to move mouse cursor to the “Start” icon located at the corner of the screen and perform a left click at first; then move the mouse cursor straight up to the “program” and turn mouse cursor right to a specified program, and finally perform a left click to execute the program. A real interaction is complex for a client in accomplishing each test task. However, the data collected from each task shows each client’s specific difficulties in a subtask by recording the subtask responses in detail.

2.3 Analysis Module

The CAT system measures the speed, accuracy and efficiency of each evaluation task. The speed is the time spent to accomplishing a single test item correctly. The accuracy is the percentage of correct responses from all trails in each test task. The efficiency is the trajectory of the cursor moving, and is measured by two indicators, the number of submovement and ratio of actual moving path and distance of a task. Speed and accuracy are defined as quantitative indicators of mouse operation performance, while efficiency is a qualitative indicator.

Since the CAT system is case-based approach, the evaluation data were recorded in each single client’s database. Rather than analyzing a group subjects’ average performance, the Analysis module is designed to understand each client's performance in a single assessment or the difference between two assessments. Figure 1 shows the quantitative information of speed and accuracy shown. The Analysis Module is adopted to analyze the cursor movement trajectory. Figure 2 indicates the trajectory originated from real cursor movement during a dragging task. The data in Figure 2 reveals that the client could move the icon to the destination, but could not release the button when the icon was there. Furthermore, as the trajectory indicated in Figure 2, the difficulties this client met were :1) starting to move the icon toward to destination, 2) targeting on the destined area correctly, and 3) releasing the icon inside the destined area.

Additionally, the program records the cursor coordinates at every decisecond. The coordinate data demonstrate the pattern and fluency of movements. Figures 3 and 4 indicate the speed changed during a task of "targeting and left-single clicking". In the figures, the *x*-axis denotes movement time (decisecond), and the *y*-axis represents distance (pixels). The performance result reveals that the cursor was moved intermittently. The cursor speed increased quickly, and then decreased immediately. Figure 4 indicates a different scenario. The client needed about 2.5 seconds to start mouse cursor, and then created two major submovements before reaching the target. He activated the left-single click about 2.5 seconds after the cursor had been finally stopped.

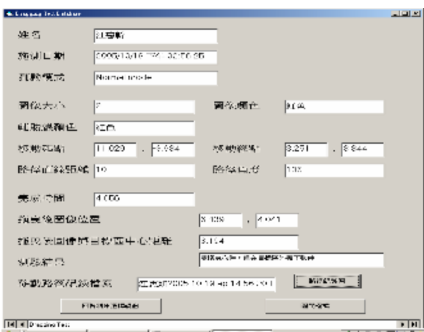


Fig. 1. An example of outcome screen of a dragging task

Fig. 2. An example of path of a dragging task

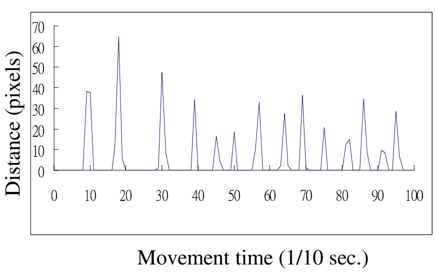


Fig. 3. An example of cursor moving with many submovements

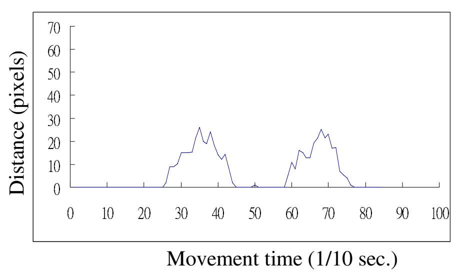


Fig. 4. An example of cursor moving with two major submovements

3 Comparability

Some comparability tests were performed with a trackball, a joystick and a multi-switch mouse, which are popularly used as alternative point and selecting devices. Two undergraduates who were familiar with mouse manipulating participated in the evaluation. The comparability test results demonstrate that these devices can complete all evaluation tasks in the CAT system.

4 Discussion and Conclusion

This study reports a case study of choosing the appropriate device for a 5^{th} grade girl with cephalic paralysis (CP). This girl could not interact with a computer through a standard mouse, but she needed to read independently and effectively on a computer. To select an appropriate device, two kinds of multi-switches mouse, Star (a switch mouse manufactured by TASH Inc.) and grouped switches mouse (five single switches) equipped with a Mouse Mover (an adaptor, also manufactured by TASH Inc.) were selected as candidates to examine their performance by conducting a training program after the revealing performance of five devices. Figure 5 illustrates the performance result of the two devices in “targeting and clicking”. As Figure 5 indicates, by using Start, she could moving horizontally and vertically well, but her diagonal moving performance was poor even after practicing eleven times. However, she performed very well when using the grouped switches mouse, even when the target square size was reduced from $0.64cm^2$ to $0.16cm^2$. Finally, the grouped switches mouse was adopted to assist her to learn to read digital text on the computer.

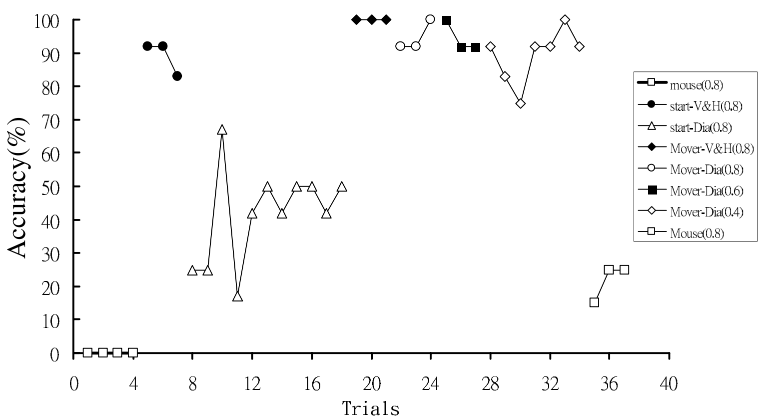


Fig. 5. Accuracy of cursor moving using three devices

5 Discussion and Conclusion

This study describes the major tests of a computerized assessment tool and its application for selecting proper pointing and selecting device for a client. The CAT system was developed through an extensive literature review, task analysis of fundamental mouse operations and expert review. Clinical rehabilitation professionals can adopt the CAT system to help them to evaluate a client's computer interaction performance. The CAT system comprises three subsystems, which not only involve the full range of needed mouse tasks, but also the essential functional mouse activities when interacting with a computer. The preliminary application of the CAT system indicates that the CAT system is useful in clinic intervention. However, further clinical application studies involving individuals with different disabilities or limitations are still needed in the future. Additionally, an expert system to diagnose the most appropriate) operating environment and pointing devices may be integrated into the future system to provide clinicians with the most valuable data.

Acknowledgment

The authors would like to thank the National Science Council of the Republic of China for financially supporting this research under Contract No. NSC 94-2524-S-415-001-.

References

1. Alliance for Technology Access: Computer and web resources for people with disabilities (5th ed.). Alameda, CA: Hunter House (2004).
2. Anson, D. K.: Alternative Computer Access: A Guide to Selection. F. A. Davis, Philadelphia (1997).
3. Anson, D., Lawler, G., Kissinger, A., Timko, M., Tuminski, J., & Drew, B. Efficacy of three head-point devices for a mouse emulation task. Assisstive Technology, (2002)14(2), 140-150.
4. Chen, M. C., Meng, L. F., Hsieh, T. F., Chu, C. N., & Li, T. Y.: Computerized assessment tool for mouse operating proficiency Lecture Notes in Computer Science, (2004)3118, 849-856.
5. Capilouto, G. J. :Movement variability and speed of performance using a head-operated device and expand membrane cursor keys. Lecture Notes in Computer Science, (2004)3118, 820-826.
6. Hwang, F., Keates, S., Langdon, P., & Clarkson, J.: A submovement analysis of cursor trajectories. Behavior & Information Technology, (2005)24(3), 205-217.
7. Kadouche, R., Abdulrazak, B., & Mokhtari, M.: Designing an evaluation method for computer accessibility for people with severe disabilities. Lecture Notes in Computer Science, (2004) 3118, 845-848.
8. Lane. A., & Ziviani. J.: Introduction to the Test of Mouse Proficiency. The Occupational Therapy Journal of Research: Occupation, Participation and Health, (2002)22(3), 111-118.
9. Mazer, B., Dumont, C., & Vincent, C.: Validation of the Assessment of Computer Task Performance for Children. Technology and Disability, (2003)15, 35-43.
10. Wright, C., & Nomura, M.: From Toys to Computer: Access for the Physically Disabled Child. Authors, San Jose (1985).
11. Wu, T. F., Meng, L. F., Wang, H.. P., Wu, W.T., & Li, T. Y.: Computer Access Assessment for Persons with Physical Disabilities: A Guide to Assistive Technology Interventions. Lecture Notes in Computer Science, (2002)2398, 204-211.

People with Disabilities: Assistive Homes and Environments

Introduction to the Special Thematic Session

Gerhard Nussbaum

Kompetenznetzwerk Informationstechnologie zur Förderung der Integration von Menschen mit Behinderungen (KI-I), Hauptstraße 119, 4232 Hagenberg i.M., Austria
Gerhard.Nussbaum@ki-i.at
http://www.ki-i.at

Abstract. Assistive Homes and Environments are a category of Smart Homes and Environments and are an approach to independent living. All important devices are linked together and allow their integrated control by an accessible user interface. These environments have the ability to compensate some of the occupant's disabilities and therefore enhance the occupant's independence. This special thematic session deals with Assistive Homes and Environments and related developments and research. The topics reach from "Smart and Assistive Homes" over "Control of ICT devices by eye gaze" to "Health monitoring systems for elderly and disabled people".

1 Introduction

Nowadays the technological revolution moves more and more into our direct environment. Houses, environments and things of the daily living are getting smarter and smarter. Ambient, pervasive and ubiquitous computing are concepts and often heard words in science. "Everything gets smart" is one of the slogans of ubiquitous computing..

In [1] Mark Weiser, the pioneer of pervasive computing envisioned "a physical world that is richly and invisibly interwoven with sensors, actuators, displays, and computational elements, embedded seamlessly in the everyday objects of our lives, and connected through a continuous network." This vision is a basic concept of Smart Environments.

Smart houses, homes and environments are nowadays no longer science fiction. Several smart houses already exist in different countries all over the world as research objects and the idea of Smart Houses is becoming more and more common. There are several definitions and descriptions of Smart Homes and Environments and some are as follows:

"A small world where all kinds of smart devices are continuously working to make inhabitants' lives more comfortable" is the description of Smart Environments in [2].

"A 'smart house' is one whose subsystems, such as security, lighting, entertainment, heating, cooling and communications, work together enhancing each other's capabilities. A smart house uses computers to help its occupant's live healthy,

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 457 – 460, 2006.

happy and safe lives by performing many tasks automatically to help manage the household." is the description of a Smart House in [3].

"A smart home or building is a home or building, usually a new one that is equipped with special structured wiring to enable occupants to remotely control or program an array of automated home electronic devices by entering a single command." is the description of a Smart Home in [4] (Fig. 1).

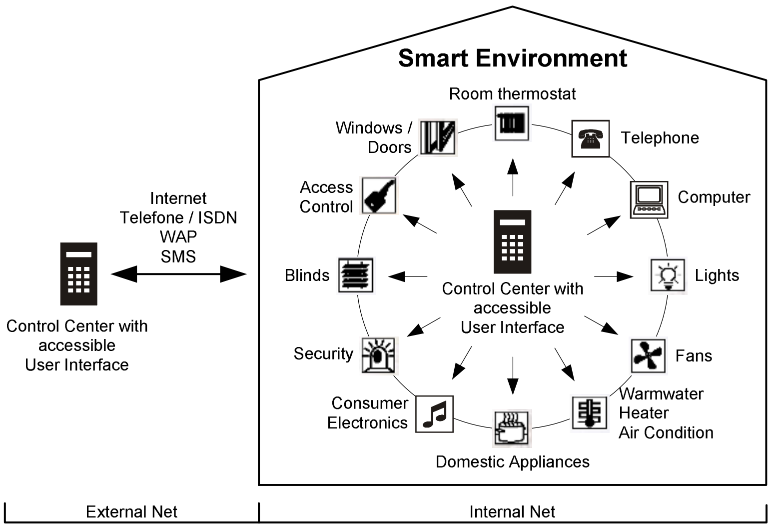


Fig. 1. Schematic and functional depiction of a Smart Environment [5]

One aim of smart houses is to enhance the comfort of the occupant. The occupant e.g. can control his house and all devices like consumer electronics, domestic appliances and other electric devices from his couch. He can check if the washing machine has finished its programme [6], he can turn on the oven [6], he can open or close all blinds of the house with just a finger tip [7], he can control the air condition [6], he can activate a special light scenario for a romantic mood [7] and while he is watching TV the automatic vacuum cleaner is cleaning the room [8]. These are only some examples which can be realised with off-the-shelf home automation systems and domestic appliances. Smart houses have much more functionalities and features.

2 Assistive Homes and Environments

All this comfort can be an essential benefit for occupants with disabilities. Such equipped environments have the ability to compensate some disabilities of the occupant [9]. It can simplify their daily lives and reduce the dependency on other persons. They can use all the provided functionality of the environment by themselves, reinforce their independence and personal freedom and can therefore longer remain in their usual surroundings. The environment itself can become assistive if the user interface is accessible to the occupant with disabilities.

Unfortunately such smart homes are currently far too expensive for persons living in social housing. Unlike in smart homes the features and functionalities are limited to the ones the occupant with disabilities benefits from which makes the realisation cheaper.

In [10] assistive homes are defined as follows: “The Assistive Home is an environment in which ‘all’ electrical items and important appliances are linked together and which allows the user the integrated control of them. Besides the user interfaces of the items and appliances there is also a central control centre with an accessible user interface based on standard HCI components.” (Fig. 2) To control the Assistive Environment off-the-shelf environmental control devices like FST James 2000 [11], SiCare devices [12], Gewa Prog 3 [13], etc. also can only be used with limitations: these devices only support unidirectional communication. They cannot receive state events and states from devices and present them to the user.

Not linkable devices like consumer electronic devices, telephones, etc. can be controlled directly via IR commands by the control centre or environmental control device.

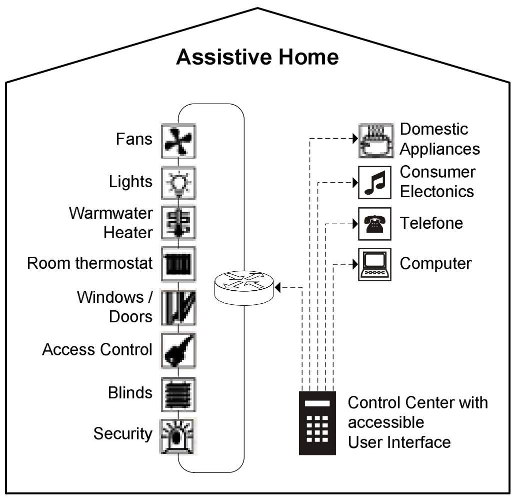


Fig. 2. Schematic and functional depiction of an Assistive Home [10]

To enhance the security of occupants with disabilities systems like fall and motion detectors, health monitoring, fire and smoke detectors, etc. also can be included if it is required. Some of these systems – especially health monitoring and eHealth applications – are topic of current research and development.

3 Conclusion

Assistive Homes and Environments can extend disabled occupant’s independence and personal freedom and therefore improve their quality of life. Such environments are no longer science fiction.

This special thematic session takes a look at current research and development in the area of Assistive Homes, Smart Homes, Health Monitoring, Communication Devices, ICT device control by eye gaze, etc. Aim of this session is to exchange knowledge and experiences and to push forward the research and development related to Assistive Homes and Environments.

References

1. Weiser, M.: The Computer for the 21st century. Scientific American, Vol. 265, September 1991, p. 66-75
2. Cook, D., Das, S.: Smart Environments: Technology, Protocols and Application. Wiley-Interscience, November 2004
3. Frost, T.F.: The Everyday Life of a Household in Cyberspace. Open Session / Workshop "Quality of life facing the 21st century - How standards will help", 1999, Beijing, China, ISO, http://www.iso.org/iso/en/commcentre/presentations/ga/gaopen/1999/ga1999-03.pdf
4. Redriksson, V.: Smart Home or Building. A Whatis Definition. http://whatis.techtarget.com. 2001
5. Aebischer, B., Huser, A.: Vernetzung im Haushalt – Auswirkungen auf den Stromverbrauch. Swiss Federal Office of Energy. November 2000
6. Siemens: serve@home. http://www.servehome.com/
7. Konnex Association: Mir ging sofort ein Licht auf: EIB/KNX ist intelligente Systemtechnik für Wohnkomfort, Sicherheit und Wirtschaftlichkeit. Konnex Brochure, 2004
8. Electrolux: Trilobite 2.0. http://www.electrolux.at/node142.asp
9. Dewsbury, G.A., Edge, H.M.: Designing the Home to meet the Needs of Tomorrow … Today: Smart technology, health and well-being. Open House International, Vol26, No2, pp33-42. 2001
10. Nussbaum, G., Miesenberger, K.: The Assistive Home – More than Just Another Approach to Independent Living?. Lecture Notes in Computer Science, Volume 3118, Jan 2004, Springer-Verlag Berlin / Heidelberg, Pages 891 – 897
11. FST – Fondation Suisse Pour Les Téléthéses: Selbständig Leben… dank James 2. http://www.fst.ch/FST2/CATALOGUE/fiches_produits_details/James2_D.pdf
12. Dr. Hein GmbH: SiCare Series. http://www.sicare.de/sicare_produkte.html
13. GEWA: GEWA Prog 3. http://www.gewa.se/english/products/prodfra2.html

Prototyping and Evaluation of New Remote Controls for People with Visual Impairment

Michiaki Yasumura, Ryo Yoshida, and Megumi Yoshida

Interaction Design Lab., Keio University at SFC
Fujisawa Kanagawa 252-8520, Japan
yasumura@sfc.keio.ac.jp
http://web.sfc.keio.ac.jp/~yasumura

Abstract. Many new home appliances being introduced to the market cannot be fully used by people with visual impairment due to lack of non-visual feedback. We have developed two new types of Remote Control(RC)'s for people with visual impairment, one with voice recognition and the other with numeric keys. We conducted two types of evaluations for them. We found that while the key RC took more time to use than the voice RC, the participants were more satisfied with the key RC.

1 Introduction

With increasing research interest in ubiquitous computing, smart homes are being built as test beds. New types of home appliance using home networks are becoming available as commercial products. However, usage of these sophisticated new home appliances is still low for people with disability, and especially for people with visual impairment. The new appliances tend to use flat buttons and LCD displays that provide visual feedback only, or only a little sound/ tactile feedback.

We studied how home electronics were used by people with visual impairment; initially with a detailed interview with two participants and subsequently with a questionnaire administered to 20 respondents. We found that people with visual impairment typically use only a small part of what the appliances were capable of, and that they wanted to have more non-visual feedback.

We then developed two types of air conditioner remote control (RC) for people with visual impairment; one type used voice recognition and the other used key input (ten numeric keys). These RC's were then evaluated by two samples of participants; one containing visually un-impaired people and the other containing people with visual impairment.

2 Related Work

The researches on the ubiquitous computing are active now in world wide, and the one of the hot issues is ubiquitous computing at home [14,6]. Among the many such research activities, there are two trends; one is the ethnographical

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 461–468, 2006.

ersearches on ubiquitous home [7,8] and the other is the development of smart home appliances [9,10,11,12,13]. Most of the target users of these smart home appliances seem to be un-impaired person, at most elderly person, but not person with visual impairment, since they often use LCD and other displays.

There are some researches on remote controls for person with disability; one of the most famous one is the Universal Remote Console (URC) by Trace R&D center [2] and another one is the Total Access System (TAS) by CSLI [3]. Both of them are very good scheme for universal RC, but they are only general one as well. We would like to know more concrete problems of remote controls for people with visual impairment.

We had developed a remote control with voice recognition [5], and some companies made voice operated RC's as commercial products [16]. These systems lack system feedback; they could issue commands but there were no ways to know the status of appliances.

3 The Design of New Remote Controls

Before the developments of new remote controls, we conducted the hearing with two person with visual impairment on their use of home appliances. They used their home appliances with the restricted functions that they could memorize. They wanted to have more feedback from them. They also claimed that the memory load was not negligible because for example, they forgot the layout and the function of RC buttons for air conditioners.

Then we sent out questionaires on home appliances and received 20 responses from people with visual impairment. We found that most difficult appliance for them to use was air conditioner. For the RC operation, most of them recognized the buttons by position. The most serious problem they felt on RC was the lack of feedback. Finally, we asked them whether they wanted to use voice recognition or not. Almost all answered *YES*.

Therefore we decided to developed two new remote controls; one used voice recognition, the other used ten numeric key pad. We thought that key pad had tactile feedback that blind person might feel easy to use. Both RC's should have voice feedback of the status of appliance. Thus the target home appliance should be information appliance that is connected with network. We selected air conditioner as the first target appliance.

3.1 Description of the Remote Controls

We developed two types of RC; one used voice recognition (the voice RC) and the other used numeric keys (the key RC) for the air conditioner (AC) connected to a network through Bluetooth. The common functions of two RC's are as follows.

- Power on and off ; check the power status; when the power is on, it gets the current temperature and humidity.
- Set the operation mode ($Auto > Cool > Warm > Dehumid > Wind$) ; check the operation mode
- Set the temperature ; check the temperature

- Set the humidity ($Low(40\%) > Middle(50\%) > High(60\%)$) ; check the humidity
- Set the wind flow ($Slight > Weak > Strong > Auto$) ; check the wind flow

The function of "check" is to show the status or the current value of each function. The user knows these values or status by looking at the LCD display with this check function.

3.2 The Voice RC

The Voice RC is a device that controls home appliance by recognizing the human voice. The control by voice command is based on natural language and therfore quite intuitive because it can utilize common verbalizations. Thus our early surveys revealed that the voice recognition might fit well for practical use. We had already developed a voice command system, but the devloped system was one-directional infra-red RC. In this study we used an air conditioner (AC) conected to a network. The status of AC, such as, temperature, humidity, and other could be retrieved through network. We developed a voice RC system that provides several basic commands to set and to check the status of AC. The control part of the system was developed by using MS C# that utilizes the voice recognition engine called Julius [17] for SAPI and the voice synthesis program called MS SAPI [18].

3.3 The Key RC

The key RC is a device with numeric key input and uses synthesized text-to-speech feedback. We used the USB type ten keys with five rows four columns buttons to control the network air conditioner(AC). We have developed the key RC system with C# to accept the key input and to control MS SAPI [18].

We devised the allocation of buttons considering the regularity of keys' usage to alleviate user's memory load. For example, we assigned the big right lower button to switch on or off the power. The left column of the buttons was for setting and the right column was for checking. The middle column was for the temperature control; the upper middle was for raising and the lower middle was for lowering the temperature. The center button with small boss is the anchoring point and used for help button (Figure 1).

For the key RC, we developed a new feature called "Help mode" that provided the user the meaning of each key. By pressing the help button, the system switches to the help mode. In the help mode, the pressing of any key except of the help key yielded to a voice output telling the function of the key. The help key was located in the center of the numeric key in order to reduce the memory load.

Every key operation accompanies the voice feedback. Some of the examples of voice feedback are as follows.

- For the temperature setting: "The room temperature was fixed at 27 degree Celsius."
- For the wind-flow checking: "The wind flow is strong now."

Fig. 1. The layout of buttons for key RC

- For the switching to the help mode: “Now we switched to the help mode.”
- For pressing the operation mode button in the help mode: “This button will change the operation mode. It changes the mode in the order of Auto, Cooling, Warming, dehumidifying, blowing.”

4 Evaluation

The evaluation was carried out by a total of 19 participants. Eleven of the participants were normal sighted but wore an eye-mask to simulate loss of sight. The other eight participants were blind. Each participant in the study evaluated three RC’s: the original (normal) RC that came with the air conditioner, and the voice RC and the key RC’s developed for this study.

Due to the space limitation, we skip the description of the first evaluation test that was for participants with normal sight. The result was almost similar to the second test that was for blind people.

For the experiment of the normal RC, at first we explained the configuration and the function of the buttons we would use for the experiment that the alarm would come out when the setting was changed, and the rough configuration of the AC to the participant. The normal RC contains the power button and the operation mode button. They are located in a semicircle. By opening the slide cover, the wind flow button, and others appeared.

We explained the voice commands that would be used for the voice RC to the participants. For the key RC, we explained the configuration and the function of keys, and the help mode to the participants.

5 The Evaluation Test for Blind

5.1 The Method

We conducted the experiment for the person with visual impairment. The number of participants was 8 in total; 7 male and 1 female. All participants were blind. The initial setting for the experiments was as follows:

$Power : off; Operationmode : Auto; Temperature : 25degree; Windflow : Auto;$

The requested tasks are as follows.

- Task1: Power On
- Task2: Check the operation mode
- Task3: Change the operation mode to Cool
- Task4: Check the current room temperature
- Task5: Raise the room temperature by two
- Task6: Change the wind flow to Strong

5.2 The Test Results

We measured the elapsed time for each RC. Figure 2 shows the average time needed for task completion. In task 2&4, the task times for the normal RC were not available, and were omitted. Table 1 shows the sum of completed task time for each RC. Table 2 shows the number of incomplete tasks for each RC.

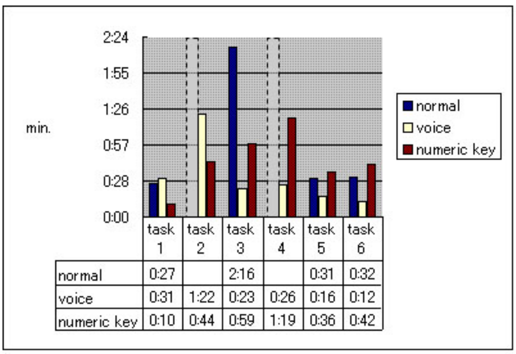


Fig. 2. The Evaluation Test Result (Task Time)

Table 1. The Sum of Completed Task Times for each RC

The normal RC	The voice RC	The key RC
4:10	3:46	4:26

Table 2. The Number of Incomplete Tasks

The normal RC	The voice RC	The key RC
18	6	1

The task time of the voice RC in this experiment was longer than the first experiment, but it was the shortest among the three RC's[1].

[1] We excluded the task times of task 2&4 for the normal RC, because they were unable to do it. Therefore, actually they took infinite time.

5.3 Interviews

We received the following comments by the participants afterward.

The Normal RC. First of all, most of the participants claimed that it was difficult to know the distance and the direction to the AC, in spite of the explanation of the layout of the AC at the starting point of the experiment. This is the big issue since the RC has to be correctly directed toward the appliance.

Also many participants claimed that they knew the sound of power on and off, but they requested the feedback sound of full circle point for the toggle button. The current system returns the sound of "Pipi" for the power on and "Piee" for power off, but for the toggle button it always returns the same sound "Pi". They requested the unique sound when the toggle switch goes around to the original point. Some of them also requested to assign the different shape for the different button. Some said that the feedback sound was too small to hear.

The Voice RC. Some of the cases, the voice recognition didn't work well. Therefore most of the participants said it was alright when the voice recognition worked well, but it would be useless if the voice recognition failed.

As for the voice command, some of the participants said that they preferred the shorter commands, and that it was difficult to remember the command.

Other opinions were as follows.

- The error handling was difficult.
- The voice feedback was sometime frustrating.
- The voice RC was not casual, tough it worked smoothly.
- The operation was easy to understand.

The Key RC. There were many positive opinions for the key RC as follows.

- It was easy to use after getting accustomed to it.
- The users could be confident with it because numeric key was familiar device in vending machines or in ATM's.
- The memory load was small because it had help mode.

Most of the participants claimed the voice feedback problems. In particular, for the quick and succeeding push's for the temperature set, the voice feedback was too slow, and it required the short, quick responses.

General Comments. The general comments for all RC's were as follows.

- The current RC's have too many functions.
- The operation mode switch will be used twice in a year, before summer and before winter. One will easily forget the layout of the buttons.
- The voice feedback should be minimum.
- The best thing is to have voice feedback for the normal RC.

5.4 The Overall Evaluation

We asked the participants to evaluate each RC in 5 levels whether he or she would like to use the RC in the future. The result was in average the key RC was 4.4, the voice RC was 2.9, and the normal RC was 2.4 (Table 3).

Table 3. The Overall Evaluation of each RC

The normal RC	The voice RC	The key RC
2.4	4.4	2.9

6 Discussion

It may be possible to design home appliances in smart homes or assistive homes without help or interactive directions. However, even when a home or its appliances are smart enough to know the context and the residents' preferences, users may sometimes want to override automated settings and input control settings themselves.Thus simple, yet effective RC's will likely remain necessities, even for smart appliances in smart homes.

Although the voice RC was less preferred than the key RC, it led to faster performance, and its acceptability may increase with further design enhancement, particularly through integration of voice control with some form of agent [1]. The authors believe that voice interaction could be more useful with some agent. In this view, future voice RC's may serve not only as command interpreters, but also as peers that chat or negotiate with the user. For example an agent may advise the user on power load or power consumption. Improved help function should also lead to better acceptance and use of RC's.

7 Conclusion

We developed two new Remote Controls for people with visual impairment to control an air conditioner. The results reported above show that the RC's are effective and useful. Useful directions for future research include adding help functions for the voice RC, and verifying the universality of our RC by adapting it to other appliances.

References

1. Schneiderman, B. and Maes, P. : Direct Manipulation vs. Interface Agents. *ACM Interactions*, Vol. IV.6, (1977) 42–61.
2. Zimmermann, G., Vanderheiden, G. and Gilman, A. : Prototype Implementations for a Universal Remote Console Specification. *CHI '02* (2002) 510-511.
3. Scott, N. and Galan, J. : The Total Access System. *CHI '01* (2001) 13-14.
4. Tsukada, K. and Yasumura, M. : Ubi-Finger: Gesture Input Device for Mobile Use. *APCHI '02*, (2002) 388–400.

5. Yasumura, M., Iga, S., and Fukui, T. : The prototyping and the evaluation of a voice interaction system, *Keio University, Research Memo*. KEIO-IEI-RM-93-009, (1993). [in Japanese]
6. Edwards, W. and Grinter, R. : At home with ubiquitous computing: seven challenges. *UbiComp '01*, (2001) 256-272.
7. Crabtree, A., Rodden, T., Hemmings, T. and Benford S. : Finding a Place for UbiComp in the Home. *UbiComp '03*, (2003) 208-226.
8. Morris, M., Lundell, J., Dishman, E. and Needham, B. : New Perspectives on Ubiquitous Computing from Ethnographic Study of Elders with Cognitive Decline. *UbiComp '03*, (2003) 227-242.
9. Humble, J., Crabtree, A., Hemmings, T., Åesson, K., Koleva, B., Rodden, T. and Hansson, P. : "Playing with the Bits" User-Configuration of Ubiquitous Domestic Environments. *UbiComp '03*, (2003) 256-263.
10. Tarrini, L., Bandinelli, R. and Miori, V. and Bertini, G. : Remote Control of Home Automation Systems with Mobile Devices. *Mobile HCI '02*, (2002) 364-368.
11. Tapia, E., Intille, I. and Larson, K. : Activity recognition in the home setting using simple and ubiquitous sensors. *PERVASIVE '04*, (2004) 158-175.
12. Green, W., Gyi, D., Kalawsky, R. and Atkins, D. : Capturing user requirements for an integrated home environment. *NordicCHI '04* (2004) 255-258.
13. Saito, A., Minami, M., Kawahara, Y., Morikawa, H. and Aoyama T. : Smart Baton System:. A Universal Remote Control System in Ubiquitous Computing Environment. *ICCE '03* (2003) 308-309.
14. Georgia Institute of Technology : Aware Home Research Initiative. `http://www.awarehome.gatech.edu/`
15. Isteon : Smarthome. `http://www.smarthome.com/`
16. Innotec Systems Inc. : Voice operated Remote Control. `http://www.innotechsystems.com/voicefire.htm`
17. Open Source Development Network :Julius. `http://julius.sourceforge.jp/en/julius.html`
18. Microsoft : SAPI. `http://www.microsoft.com/speech/download/sdk51/`

Constructing Adaptive User Interface Through Semantic Descriptions

Hyun Namgoong, Kyung-il Kim, and Yun Koo Chung

Computer Software & Engineering, University of Science and Technology
52, Uheun-Dong, Daejeon, Korea
Intelligent Robot Research Division
Electronics and Telecommunications Research Institute
161, Gajeong-Dong, Daejeon, Korea
{ngh, zbmon, ykchung}@etri.re.kr

Abstract. We present semantic descriptions for dynamically integrated usage of the devices providing user-centric services. This paper describes the architecture of the user-side controller for adaptive user interface supporting control of the devices. It provides users with easy interface for controlling heterogeneous devices in environment. Especially, it is efficient for people with disabilities to operate the surrounding devices. Furthermore, it can be applied to automated home control and proactive service system, etc.

1 Introduction

In ubiquitous computing environment, various input and output devices exist for supporting user convenient services such as automated home appliances, voice recognizers, location sensors, etc [1]. In order to use various devices with felicity, people should have knowledge how to exploit heterogeneous interfaces suitable for the devices. Since it makes them inconvenient for interacting with various devices, the needs for integrated interface are increased [2]. The integration of the devices for intuitive and easy usage requires to describeg them with normalized way such as XML based description [3].

We propose the adaptive interface generation which supports the easy usage of multiple devices. In our system, the devices are marked up with semantic description which enables to provide the logic for describing information resources with their semantics. In addition, a controller in user's mobile device finds the description from web or local spots to make adaptive interface acceptable to user.

2 Modeling of Devices

We describe semantically every device through ontology for the dynamic integration of the devices in the environment. Each device for sensing and offering services can has heterogeneous characteristics according to its middle-ware, such as access pattern, related environmental element, its functionalities, and so on. Ontology, which is the specification of a conceptualization, offers the semantics for the representation of

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 469 – 472, 2006.

those features. The normalized description through the ontology, i.e., Semantic Web, provides the way for modeling and integrating as semantic resources.

We introduce three types of ontologies for modeling of the devices. The ontology is conducted by XML based standard, e.g., OWL, etc, and contains concepts and their semantic relations [4]. Each organized ontology shows the device's characteristic, such as, related environmental elements and access information containing what the device is, what environment is related with and how we can access to it.

2.1 Concept Ontologies

Three concept ontologies are presented in our prototype for modeling devices semantically. It supports the description for three kind of semantics; device, environment and message. Device ontology deals with general description of the device using taxonomy trees. Environment ontology contains the concept related with environmental elements. Message ontology is focused on the concept describing access information for the devices.

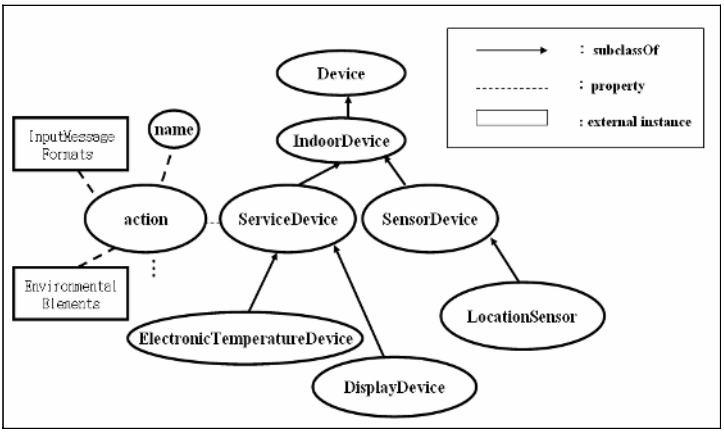


Fig. 1. Device ontology

The concept ontologies enable the devices systemically described with specific information containing its own features. Device ontology includes overall characteristics of each device such as classification and containing properties. As shown in Fig. 1, ElectronicTemperature Service is subclass of ServiceDevice, and it also has actions indicating functionality of the devices. The action has its own functional name and is expanded with the instance of *InputMessageFormats* class and the instance of *EnvironmentalElement* class as properties. Each property means how the others can access to the device and which kinds of changes are occurred through execution of the action. Moreover, Input and Output MessageFormats show information indicating communication protocol, address, ports using Message ontology. For depiction of EnvironmentalElement, Environment ontology includes feasible elements in environment.

2.2 Describing Devices with the Ontologies

The description of devices depends on references to conducted ontologies. Every device is represented as an instance of the class in the device ontology. The device instance also has action instances which describe properties of *Environment* and *Message* ontology to show its characteristics. For example, the electronic temperature controller with middleware is presented as an instance of the *ElectronicTemperatureDevice* inherited *ServiceDevice*. It also has an action instance. Therefore, the action has *InputMessageFormats* in the Message ontology and the *RelatedEnvironment* in the Environment ontology. Fig. 2 shows the part of the semantic description of the electronic temperature controller which has an action with their *InputMessageInterface* and *RelatedEnvironment*.

```
<device:ElectronicTemperatureDevice rdf:ID="LocalTemperatureDevice">
  <device:name>LocalTemparatureServiceDevice</device:name>
  <device:description> Device for Indoor temperature management
  </device:description>
 <device:action rdf:id=”action1”/>
   <action:InputMessageInterface rdf:resource="http://urcsp.etri.re.kr/concepts/
ServiceMessages.owl# Temperture-DeviceMessages "/>
   <device:RelatedEnvironment rdf:resource="http://urcsp.etri.re.kr/concepts/
ServiceRelatedEnvironment.owl#TempratureDeviceRelatedEnvironment"/>
 <device:action/>
 </device:ElectronicTemperatureDevice>
```

Fig. 2. Semantic description of the electronic temperature controller

3 Constructing Adaptive User Interface

The proposed system exploits user's mobile device as a controller for gathering semantic descriptions and offering the adaptive interface to the user. As shown in Fig. 3, it consists of four main component; Interface Engine, Reasoner, Knowledge base, and Message Manager. Message Manager finds the description of newly appeared service devices in the environment, and then delivers the description to Reasoner. Moreover, Reasoner translates concepts and descriptions into other concepts applicable to prior knowledge and more applicable compositions in the Knowledge base. Finally, Interface Engine generates the user interface, and presents to a user. It also undertakes flow of processes of the system and actually constructs the user interface. If user chooses the order by the interface, the order is passed through Interface Engine, Message Manager to the devices.

When a user with a controller enters the space containing the available devices, the controller finds usable devices from its surroundings. It is available from internet, a local server or physical tags like a RFID. As the controller finds the services in ad-hoc manner, the user utilizes particular services without spatial restrictions. For example, a breath disease patient can simply set her living room at higher moisture and temperature by the generated interface. Then, if the user takes a move to another authored room with the controller, it will dynamically find the humidifier and heater in the spot, and then it operates the device to meet her needs.

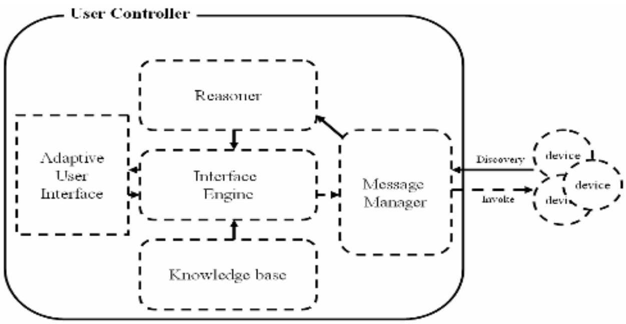


Fig. 3. Architecture of user device's controller

4 Conclusion

We presented the system for constructing adaptive user interface with semantic descriptions. In order to evaluate the proposed system, we are implementing a simulating environment. We expect that it dynamically offers user interface which make users control environment without spatial restrictions. Moreover, it enables the controller to easily generate complex and combined options with multiple functionalities of the devices using prior knowledge and user preferences. As a future work, we are planning to study federated use of the information for generating the interface. We expect that the federation can support more automated and seamless control of the devices without user's direct commands.

References

1. Wesier, M.: The computer for the 21st century, Scientific American, 265(3), pp.94-104, September 1991
2. Rekimoto, J.: A Multiple Device Approach for Supporting Whiteboard-Based Interactions, In Proceedings of ACM CHI 98 Conferences on Human Factors in Computing Systems, pp.344-351, 1998
3. Wang, X. (ed.): Semantic Space: An Infrastructure for Smart Spaces, IEEE Pervasive Computing, 3(3):32-39, July-September 2004
4. OWL.: `http://www.w3.org/TR/owl-features/`

A Proposal for a Home-Based Health Monitoring System for the Elderly or Disabled

HyungJun Kim, Bart Jarochowski, and DaeHyun Ryu

Division of Information Technology
Hansei University, Korea
harry@hansei.ac.kr

Abstract. It has been shown that the elderly, disabled, and those suffering from chronic illness benefit from early detection of symptoms and being involved with the management of their own conditions. We propose a home-based monitoring system which will continuously and unobtrusively (depending on a patient's condition and the sensors used) monitor a patient's condition. The monitoring system is supplemented by a monitoring/data warehousing service which is the first line of response in the event of an emergency. This system will be implemented using Zigbee technology.

1 Introduction

The improvement of medical technologies, techniques, and drugs in developed countries has resulted in a steady increase in the life expectancy of people in these countries. This, combined with the lower birth-rates seen in many of these developed countries has resulted in a significant growth in elderly populations. Given that improved medicine and medical technologies are largely responsible for the increased life expectancy of people, it is of no surprise that the elderly, disabled, and those suffering from chronic conditions increasingly rely on these technologies and medicines in order to maintain their quality of life. Maintaining quality of life standard, however, does place some financial burden on individuals and government. In the United States, an estimated 8 million of Medicare's 44-million beneficiaries with five or more chronic conditions account for about 66 percent of Medicare spending, according to a 2002 study by Partnership for Solutions, a policy research program at Johns Hopkins University financed by the Robert Wood Johnson Foundation. Given that nearly 40 percent of elderly people have some sort of chronic condition that limits function and causes morbidity and 70 percent of all deaths can be attributed to some kind of chronic disease [1], solutions that improve the lives of people with chronic conditions are urgently needed. Because of this, the focus of a lot of recent research has been to find ways of maintaining a high standard of care for the elderly and those suffering from chronic conditions while reducing costs and preserving or improving the standard of living that those patients have become accustomed to [2,3]. One solution is to incorporate a ubiquitous health system in the home. It has been shown

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 473–479, 2006.

that a ubiquitous health-care systems improve the productivity of healthcare professionals as well as facilitate a wider range of health-care services [4].

In order to maintain a patient's standard of living while allowing for a high standard of care, we propose a monitoring system that unobtrusively and constantly monitors a patient's condition. It has been shown that patients who are suffering from chronic conditions fare better when they have some degree of responsibility for their own self-assessment than those who have limited involvement in their own care [5]. Therefore, the aim is to both empower the patient allowing them to take some responsibility for ongoing self-assessment while providing doctors and/or emergency response professionals timely and accurate information relevant to the patient's condition. The proposed system will be both easy to use and maintain and will make use of common components wherever possible in order to keep costs down. The system will consist of a main server which will act as a *control station* and will have a secure connection to the internet. The server will interface with a ubiquitous sensor network based around the Zigbee system and also be connected to a monitoring service through the internet connection. This monitoring service will manage data coming from the server as well as act as a first line of response if an emergency situation arises.

The following pages will contain information regarding the user interface, sensors, topology and tech crap, usage models, and will finish with a conclusion.

2 User Interface

The user interface for our proposed system will be extremely intuitive in order to be accessible to even the most technologically inexperienced patients. The central point of access for our system will be a server with a large main liquid crystal digital (LCD) panel with touch-screen capabilities. This interface will feature large, easy-to-read lettering and will have a comprehensive and easy-to-understand help system. This panel will serve the following functions:

- *Interactive Examinations*: Utilizing streaming-video technologies, doctors will be able to interact with and, through this interaction, be able to obtain a rough assessment of the patient's psychological and physical condition. If required, the system will also be able to administer questionnaires and relay the answers to an assigned medical professional. This will allow patients whose mobility is restricted to see a doctor without making the arduous journey to the doctor's office.
- *Reminders*: The large screen will be used to remind patients to perform certain actions such as taking medication, perform exercises, etc. These reminders will ensure that patients consistently follow the regimen recommended to them by their doctor.
- *Status Reports*: Patients will, at any time, be able to examine the reading from various sensors in order to better assess their own condition, therefore empowering them to make better decisions about how to self-manage their illness. This will be displayed in an easy-to-interpret manner (utilizing color-coding and large lettering).

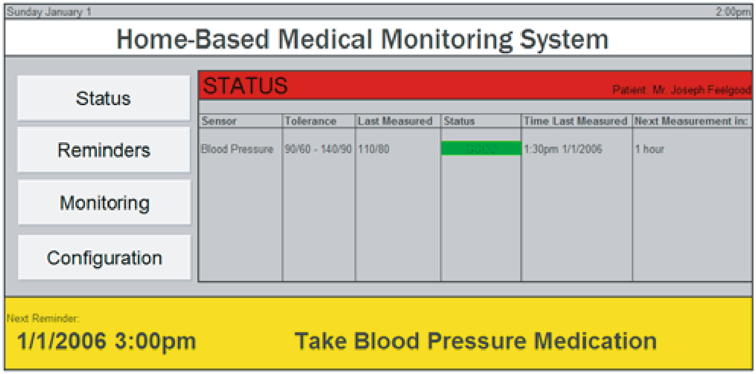


Fig. 1. An example of a user interface

– *Voice Activated Panic Sensor*: The server will feature a voice activated system whereby if the patients feel distress or falls, they can immediately get in touch with a medical professional. This will add another measure of security to the patient's life.
– *Configuration/Testing/Maintenance*: Having a central point of access will allow technicians to easily configure the system to suit its particular application.

There is also the possibility of interfacing with other household appliances and/or personal items such as a television or smart wristwatch in order to provide a more intuitive interface. An example of a user interface is shown in Fig. 1.

3 Sensors

Unobtrusively and constantly monitoring a patient's vital signs is a key challenge faced with researchers today. It has been shown that when chronic illnesses are monitored effectively, it improves the chances of a life which is both longer and of higher quality [6]. Even when a patient does not have a chronic condition, constant monitoring may be able to detect early signs of certain diseases. Evidence shows that early detection of diseases reduces the mortality and morbidity of many diseases [1].

In order to facilitate constant monitoring of patients, we will make use of a ubiquitous sensor network that is based around the Zigbee system. Zigbee has several advantages: low power consumption, data security, quick wake-up time, support for various topologies, adequate data rates, and the ability to interface with a wide variety of sensors [7].

The frequency of monitoring will depend on the sensors used as well as the patient's condition. Some patients may only require monitoring once every day (or perhaps at night), while some patients may require constant monitoring. The system will be able to be configured based on the patient's needs.

The sensors used in this monitoring system can vary according to the specific application. Sensors specific to particular medical conditions can be added as needed. There are many ways to incorporate various kinds of sensors in order for them to be as unobtrusive as possible. Since Zigbee can interface with a variety of sensors by way of an ADC (analogue to digital converter), this system is highly extensible. Some sensor possibilities are listed below:

– Toilet Sensor (measuring weight, fat, blood pressure, heart beat, urine sugar, albumin and blood in urine) [8]
– Wrist Sensor (measuring pulse)
– Bedside breathing sensor (measuring the regularity of respiration)
– Under-bedsheet sensor (measuring pulse, EKG, body temperature)
– Life-Vest [9]
– Sound sensor (to respond to calls for help)
– Location and Orientation sensor (fall sensor)

Other than sensors measuring a patient's vital signs, sensors that keep track of location and orientation may be added to automatically determine if a patient has fallen and to notify a medical response team if this has occurred. This is important because for people ages 65 and older, falls are the leading cause of injury or death [10].

Sensor configuration will be programmed at the main base station. Sensor tolerances will also be able to be configured. If a certain metric exceeds acceptable tolerances, the server will then alert the monitoring service, which will respond appropriately.

4 Data Interface

4.1 Sensor to Server: Zigbee

While there are many standards that provide wireless connectivity, Zigbee is the ideal choice for this application [7]. Zigbee is an open standard based on IEEE 802.15.4 and incorporates ad-hoc networking/routing, security, low power consumption (up to years), and is able to interface with many different types of sensors [11].

The main alternatives to Zigbee are Bluetooth and Wireless Fidelity (WiFi). Zigbee's range is greater than that of Bluetooth, which only has a 10 meter range. The high bandwidth (and consequently complexity and high power draw) of WiFi are not needed for this application.

The following is a brief features summary of Zigbee:

– Open standard
– Ease of implementation
– Dual Physical Layer (2.4GHz, 868/915MHz)
– Data rates between 20 - 250 kb/s
– Optimized for a low duty cycle (sensors)
– Low power consumption (months - years)

– Supports star, peer-to-peer, mesh topologies
– Supports many devices and networks
– 50 meter range (between 5-500meters based on environment)
– Reliable data transfer due to fully handshake protocol
– Built-in AES-128 security

These features are ideal for the implementation of the home-based monitoring system.

4.2 Platform

The system will be implemented in Java to make use of Java's portability, security, and flexibility. The Java platform also provides a very convenient environment for interfacing with Zigbee sensors and implementing remote connectivity with diverse systems. System interface is shown in Fig. 2.

5 Typical Usage

The following is an example of the typical usage of the home-based monitoring system:

1. Patient goes to a doctor for an eligibility exam.
2. The doctor records problem areas, determines acceptable sensor tolerances accounting for activities done within the home, diet, etc.
3. A 'monitoring prescription' is made up and sent to the company managing the service.
4. A technician comes to the home to install sensors, server, and configure the system with doctor-approved tolerances.
5. An account is also set-up at a monitoring center/data warehouse.
6. Patient lives his/her life normally interfacing with the system as much or as little as the patient desires (and also dependant on the patient's condition).
7. The monitoring service is on 24 hour standby (monitoring many patients at once).
8. The monitoring service is only notified when there is a significant event. (fall, sensor reading exceeding tolerance). In such an event, the monitoring service is the first-line in responding to the situation.
9. Maintenance visits may be necessary.
10. Subsequent visits to doctor: doctor may modify tolerances, add sensors (via prescription to monitoring/data warehouse).
11. Tolerances and/or sensors can be changed according to the changing needs of the patient.
12. The patient is billed for installation, hardware, and a monitoring subscription.

In this example, it is important to note that the Home-based monitoring system is quite unobtrusive. How unobtrusive the system is largely dependant on the sensors needed for the particular patient in question. For patients with mild conditions where only a few vital signs are monitored and/or the monitoring frequency is low, the system may very well blend into the background of their day-to-day lives. This is the ultimate goal for such a system.

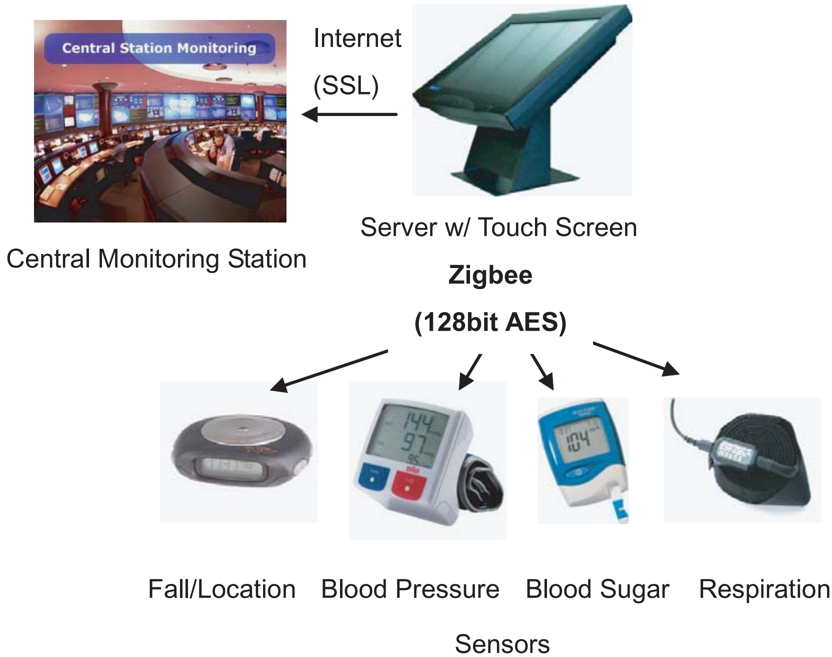


Fig. 2. System interface

6 Home-Based Monitoring vs. Remote-Monitoring (Body Area Networks)

There has been some research and development based around a body-area-network concept that is totally mobile, making use of a personal digital assistant (PDA) to provide data storage and communication to the outside world. An example of this is the EU's MobiHealth system [12]. A home-based monitoring system is very complimentary to a totally mobile system like MobiHealth. The home-based system does have a number of advantages over a completely mobile system.

A home-based system needs not necessarily include sensors that are physically attached to the body. Instead, the home-based system may perform monitoring in a more unobtrusive way by making use of sensors that interface with the body without having to be transported by the patient (for example: toilet sensor).

The home-based system also has a user interface advantage. Interfacing with a large touch-screen will be far easier for the elderly than performing similar functions on a PDA. Faced with the real limitations of eyesight and co-ordination particularly on the part of elderly patients, a PDA-based system is not as effective as the home based system (featuring a large user interface) at involving patients in the management of their own conditions.

Ultimately, the two methods are highly complimentary. In the home, there is no need for a body area network(BAN)-based system. However, outside the home, it is the only way to maintain continuous monitoring.

7 Conclusions and Future Works

The purpose of this proposal is to suggest a home-based monitoring system for the elderly, disabled, and those with chronic conditions. The goal of this system is to provide continuous and unobtrusive monitoring, while allowing the patient to live at the standard of living that they are accustomed to. This system will also allow the patient to be involved with the management of their condition by utilizing an intuitive user interface, and provide immediate response to emergency situations thereby providing the patient with a feeling of security. This system will be implemented using Zigbee due to its ideal suitability for this task. With populations aging in many developed countries, the need for a system such as the one proposed here is great. The quick implementation and deployment of such a system will undoubtedly have a positive impact both to individuals and to the entire health-care system.

Acknowledgements

This work was supported by Hansei University.

References

1. Adelman, A. W., Daly, M. P.:Geriatrics - 20 Common Problems. McGraw-Hill New York (2001)
2. Kim, Y., Kim, J.:Real-Time Health-Monitoring Network for Disabled and Elderly People. in Computers Helping People with Special Needs LNCS 3118 Springer-Verlag Berlin Heidelberg New York (2004) 377–384
3. Karshmer, J. F., Karshmer, A. I.:Computer-Based Self-health Monitoring System for the Elderly Living in a Low Income Housing Environment. in Computers Helping People with Special Needs. LNCS 3118 Springer-Verlag Berlin Heidelberg New York (2004) 385–391
4. Varshney, U.: Pervasive Healthcare. Computer (2003) 138–140
5. Lorig, K., Sobel, D. S., Sewart, A. L., et. al.:Evidence Suggesting That A Chronic Disease Management Program Can Improve Health Status While Reducing Hospitalization: A Randomized Trial. Medical Care 37 (1999) 5–14
6. Tetzlaff, L.:Consumer Informatics in Chronic Illness. Journal of the American Medical Informatics Association 5 (1998) 347–356
7. http://www.zigbee.org/en/index.asp
8. http://www.globalaging.org/health/world/toilet.htm
9. http://www.vivometrics.com/site/index.html
10. http://www.cdc.gov/ncipc/pub-res/toolkit/SummaryOfFalls.htm
11. Chevrollier, N., Golmie, N.:On the Use of Wireless Network Technologies in Healthcare Environments. in Proceedings of the fifth IEEE workshop on Applications and Services in Wireless Networks ASWN(2005) 147–152.
12. http://www.mobihealth.org/

Helping People with ICT Device Control by Eye Gaze

Fangmin Shi, Alastair Gale, and Kevin Purdy

Applied Vision Research Centre, Loughborough University
Loughborough LE11 3TU, UK
{f.shi, a.g.gale, k.j.purdy}@lboro.ac.uk

Abstract. This paper presents a computer method to help people, typically having limited mobility, to be able to operate ICT devices with eye gaze in their living/work environment. The user's eye gaze is recorded and analyzed in real-time. Any ICT device in the environment that is being looked at for a certain time period is identified, located and assumed to be the object of interest that the user wants to utilise. Through a suitable interface, the user can then decide whether to operate the device. By using this state-of-the-art technology, people with impaired mobility, or able bodied people whose movements are restricted can attain a more independent life style.

1 Introduction

Living with independence is fundamentally important to people's quality of life. However, for those who have lost significant mobility due to diseases such as ALS (Amyotrophic Lateral Sclerosis), Cerebral Palsy, spinal cord injury or Multiple Sclerosis, it is very difficult to carry out even a simple operation, for example to efficiently select buttons manually from the menu of a complex remote control.

Research in eye gaze usage can enable people with limited mobility to access technology more efficiently. Communication by Gaze Interaction (COGAIN), a European research Network of Excellence, brings together researchers from both academia and industries to research and develop eye gaze driven assistive technologies with a view to achieving effective communication and environmental control using eye gaze for mobile disabled users [1]. The research 'ART - Attention Responsive Technology' [2] described here is an example of such technology.

2 Attention Responsive Technology

The ART project's aim is to design and develop a computer-based system to locate a user's eye gaze within a three dimensional room environment and use this to identify an ICT device which they may wish to operate by a suitable interface. For example, if a user wants to switch on an electric fan, with an ART system s/he stares at the fan for a certain time period when a simple 'fan ON' option would become available for them to confirm their intention. Upon such confirmation (e.g. through a touch sensitive screen) the fan will be switched on via wireless communication.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 480 – 487, 2006.

The system currently works through two camera systems mounted on a head band worn by the user. The first camera - the eye camera- faces one of the user's eyes and tracks their eye movements, resulting in eye fixation ('gaze') data in the environment. Another compact camera – the scene camera - monitors the user's the environment in front of the user and detects any controllable devices which the user may look at. By calibrating the two cameras in advance, the position of eye gaze from the eye camera will be mapped to the image coordinate system of the scene camera. If the eye gaze is on any part of a device, the user will be offered an interface which they can then decide whether or not s/he wants to change the current status of that device. Section 3 will describe the technical details of the system. The working flow of the ART system can be seen in Figure 1.

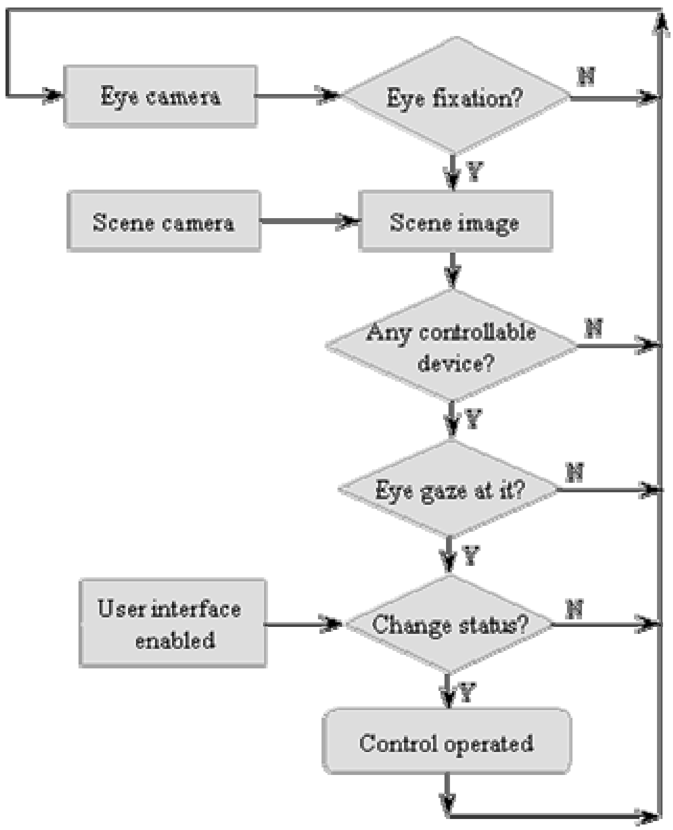


Fig. 1. ART system work flow

Compared with existing eye gaze applications, there are two main distinctive advantages of using the ART system. Firstly, its selection is relative to a real object in a three dimensional space; whereas the selections of most other systems, e.g. the Eyegaze Communication System [3], are with reference to the 2D computer screen where all devices are represented as buttons or menu items. When the number of devices increases, the layout of the screen may become more complex. The ART approach simply removes any need for a more complex menu. Second is that it eliminates any inadvertent actuation of controllable devices by introducing a simple user configurable interface for confirmation purpose. The user can therefore take the initiative to select operation of the object only when s/he wishes. Conventionally, the interaction with computers via eye movements such as Ward and MacKay [4] has an 'always-on' linking of the user's eye gaze direction to actions, which will include some un-intentional gazes. The ART method selection is therefore more objective.

3 Controlling ICT Devices with Eye Gaze

To implement the tasks as illustrated in Fig.1., a laboratory-based prototype system and its software control interface have been developed. To record a user's saccadic eye movements, a head-mounted ASL 501 eye tracker as shown in Fig.2 is used.

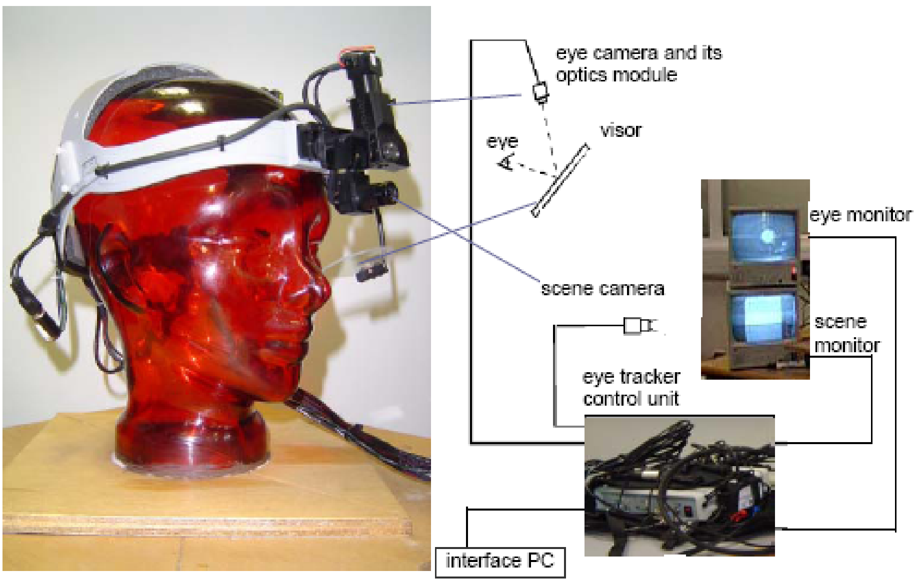


Fig. 2. ASL 501 head mounted eye tracking system

The eye, illuminated by near infra red light, is reflective from the head band visor and its image containing the eye pupil is focused onto the eye camera. The user's pupil size and the eye line of gaze with respect to his/her head, computed by the controller unit through extracting the eye pupil outline and the light reflection in the corneal, are exported as real time serial port stream data to the host PC. There is a second camera mounted on the head band which directly faces the front scene. It forms a computer vision system, together with the connected optics and frame grabber, which videos the user's front view at a rate of 50 frames per second. Two monitors display the video signal from the eye camera and scene camera respectively.

3.1 Calibration Between the User and the Two Image Coordinate Systems

By default, the eye camera records an eye pupil image of size 260x240 pixels. A user's line of gaze is extracted with respect to his/her head. Subject to the chosen adaptor to connect the direct scene camera to the PC, the size of the output scene image can be varied. The current system uses the existing PC graphics card with video input and gives an image size of 760x576 pixels. Through calibration the user's point of gaze is traced within the scene image to reflect where s/he is actually looking.

The performance of the head-mounted eye tracker system depends greatly on initial calibration. The user sits in a comfortable position while having one eye imaged on the eye monitor. A nine point target chart is placed at a suitable distance so that its image nearly fills the scene camera field of view and the user is instructed to look in turn at each of the calibration targets.

The target chart used in this project is shown in Fig.3. Its image is acquired by the scene camera within the 760x576 pixels. Through a set of image processing techniques, i.e. image noise reduction, image dilation and erosion, and edge detection, the image coordinates of the nine points can be obtained. The process and the corresponding images can be seen in Fig.3. The user is instructed to sequentially look at the nine points while the line of gaze to each individual point in the 260 x 240 coordinate system is recorded. Both sets of data are saved as input to the eye tracker controller interface. A calibration matrix is automatically generated linking the eye gaze coordinates in the eye camera system to those in the scene camera system.

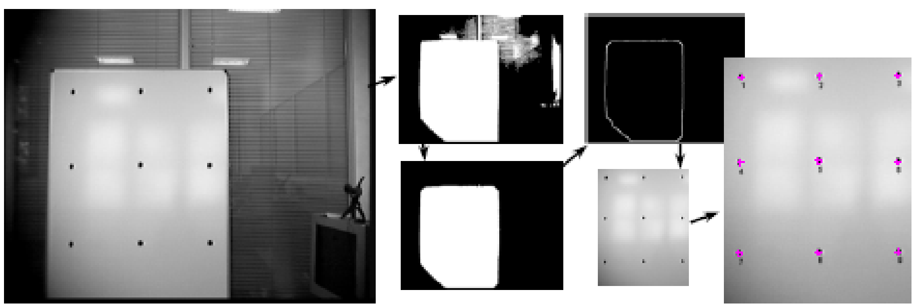


Fig. 3. Obtaining calibration target point coordinates

Typically a user calibration takes some 30 seconds. The success of such a calibration can be examined from the scene monitor where the superimposed cross hair, representing the output from the eye camera, should closely follow the real point of gaze at an object in the three dimensional room environment in real time.

3.2 Eye Data Recording and Eye Fixation Determination

In the real time application, nothing will be done until an eye fixation is determined. Simply as a starting point for development purposes, we utilise a 2s gaze at an object to indicate the user's intention to operate that object.

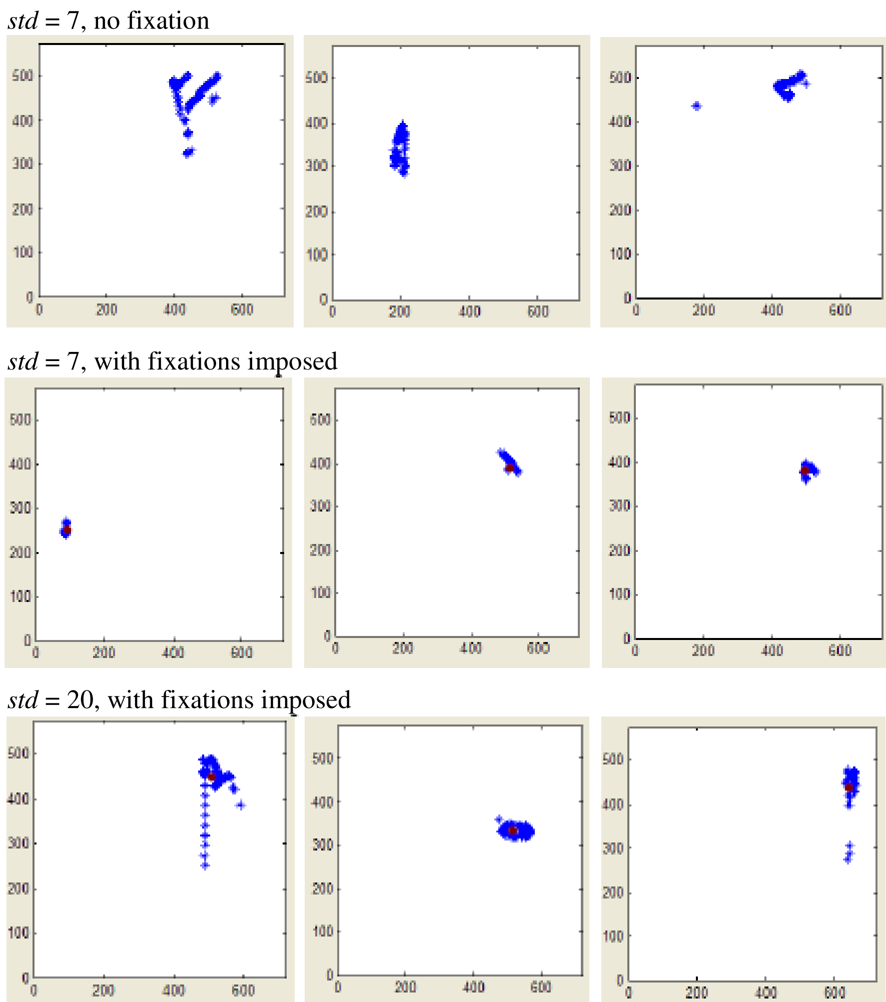


Fig. 4. Eye fixation analysis

Considering the 50Hz video frequency, 100 sets of eye data would be recorded within 2s. The eye pupil size and x, y coordinates of the points of gaze are returned, for example:

pupil size	*gaze(x)*	*gaze(y)*
46.0000	717.2308	410.4000
46.0000	717.2308	410.4000
45.0000	714.4615	412.8000

The criteria for determining an eye fixation are set as follows:

1) Remove those data where the eye pupil cannot be traced stably, either due to changes of the ambient illumination or the wrong positioning of the eye pupil image. In either case, zero pupil sizes are returned.

2) Remove those falling outside the boundary (720x576).
3) Check if the number of the remaining points is more than 60% of the original.
4) Continue to check if 80% of the left points are within a range of *mean*±2**std* (standard deviation). If yes, then a fixation is obtained.

Experiments have been made with various *std* and it has been found that a range of *std* from 7 to 20 work well to obtain a fixation from 'well' to 'loosely focused' gazes at an object. This parameter will be made adjustable in order to suit the real end users with different level of concentration abilities. Examples of the recorded eye data are given in Fig.4. It can be seen that the group data with larger *std* can tolerate a more widely spread distribution, (more tolerance to a user's head shaking or eye shifting).

A scene image may or may not contain an object of interest. It may also include more than one controllable object. SIFT (Scale Invariant Feature Transform) feature matching, developed by David Lowe [5], enables a stable approach to object identification and location. It works by saving a set of images of each pre-known object to form the reference image database, as shown in Fig 5. The SIFT feature of each image is calculated, and in real time these are compared with that of the scene image to perform image matching for object recognition. The method is described in [2] and matching constraints are discussed in [6].

Fig. 5. Example images from the reference image database

3.3 Controller Simulation Interface

The prototype system combines all of the features as illustrated in Fig.1. apart from the real action part once an effective eye fixation falling on an object is detected. Alternatively, the real time system is currently controlled by a graphic user interface for the controller simulation, as shown in Fig. 6.

By pressing the toggle Start/Stop button, the scene image can be displayed continuously in the upper right window. Upon the detection of the user's attention on a controllable device, the found object and the fixation point on it will be highlighted. The name of the objects, which are listed in the left column, will flash in different colours to invite the user's confirmation by choosing either of the radio buttons ON or OFF. Each time the gaze operation is stored in the text box 'history of eye gazes'.

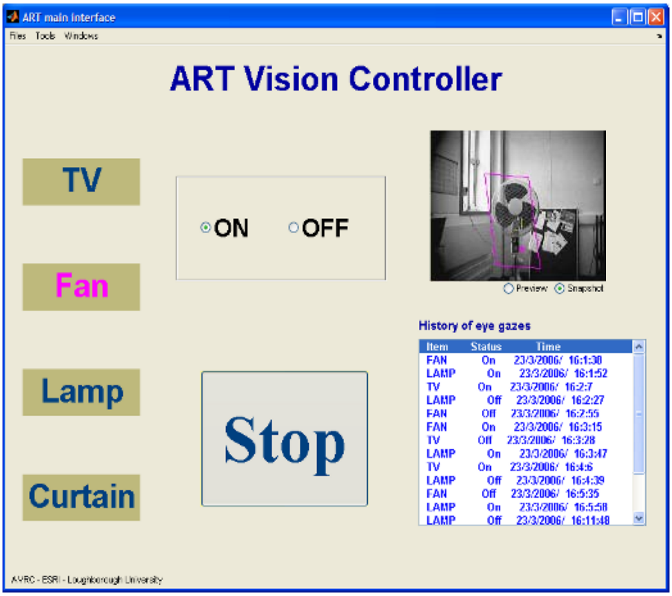


Fig. 6. A GUI to run the prototype system in real time

4 User Group

The initial research work is targeting potential end user groups and determining which objects, and object controls are most relevant to them. Additionally this work will ascertain potential user demands from such a system and the outputs from this research phase will then help shape the subsequent technical solutions formulated.

A questionnaire has been sent out to potential end users - `http://www.lboro.ac.uk/research/esri/applied-vision/projects/art/art-survey/art_q.htm`. The survey questions are composed of five main sections, including the users' general feeling of the ART concept, their requirement for the speed of using the system, the opinion on the number and the type of controllable devices, where users want to place them and finally about the users themselves. The design of the online survey takes special account of users' disability. Most of the questions are accessible by repeatedly using the Tab key and Y/N keys.

The survey is still ongoing. The initial feedback has been used as a guide for the development of the system. For instance, setting up a living room environment, the selection of controllable devices such as; TV, fan, lamp and electric curtain, and the two second item selection speed are examples reflected by the participants' answers.

5 Discussion and Future Work

The ART system currently runs on a Matlab programmed interface. The system run time can be broken into the following details: 100 eye data - 2~3s; a snap shot - 2~3s; and SIFT matching - 3~5s (40 reference images). In total it takes 7~11s. This initial slow speed can be significantly improved by either programming in C instead of Matlab or reducing the number of eye data points to be recorded.

This paper has presented a solution to the problem of using eye gaze based control systems. It offers a practical way for people with restricted mobility to make use of his/her vision ability to control devices with a view to achieving a more independent lifestyle. A laboratory-based prototype system is available with a head mounted eye tracker. A head free eye tracking system (Smart Eye) in place of the current head mounted one is under development. This will release the user from the requirement of wearing a head mounted eye movement system.

Acknowledgements

This research is funded by the ESRC PACCIT Programme. The work is also part of the COGAIN network.

References

1. Bates R., Donegan M., Istance H.O., Hansen J.P. and Rih K.-J.: Introducing COGAIN – Communication by Gaze Interaction. In: Clarkson J., Langdon P. and Robinson P. (Eds.): Designing Accessible Technology. Springer-Verlag, London (2006) 77-84
2. Shi, F., Gale, A.G. & Purdy, K.J.: Eye-centric ICT control. In: Bust P.D. & McCabe P.T. (Eds.): Contemporary Ergonomics (2006)
3. The Eyegaze Communication System: `http://www.eyegaze.com/2Products/Disability/Disabilitymain.htm`, accessed April 2006
4. Ward D.J. and MacKay D.J.C. (2002) Fast Hands Free Writing by Gaze Direction. Nature. August
5. Lowe D.G. (2004) Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision. 60, 2, 91-110
6. Shi, F., Gale, A. and Purdy, K. (2006) Sift Approach Matching Constraints for Real-Time Attention Responsive System. Proceedings of APIS5 2006, China, 31-34

IRCS – Infra Red Code System: Access to Infra Red Codes for Ambient and Assisted Living

Klaus Miesenberger[1], Gerhard Nussbaum[2], Martin Bartsch[1], Thomas Mayrhofer[1], and Christoph Strobl-Mairhofer[1]

[1] Institut "integriert studieren – integrated study", University of Linz, Altenbergerstrasse 69, A-4040 Linz, Austria
kmiesenberger@jku.at
http://www.integriert-studieren.jku.at
[2] Kompetenznetzwerk Informationstechnologie zur Förderung der Integration von Menschen mit Behinderungen (KI-I), Hauptstraße 119, 4232 Hagenberg i.M., Austria
gerhard.nussbaum@ki-i.at
http://www.ki-i.at

Abstract. This paper presents a prototype of a web based system which allows to upload and to share infra red codes for consumer electronic devices support or automate the integration of the interface of these devices in an accessible user interface i.e. in Environmental Control Systems.

1 Introduction

Environmental Control Systems (ECS), Smart Environments or Ambient and Assisted Living (AAL) Systems provide and integrate interfaces to a number of devices which people with disabilities are not able to access or use independently. Infra Red (IR) Codes are an important possibility in this field as it is intensively used for remote control units of consumer electronic devices.

Traditional ECS have to learn and store these codes for each new device and integrate it to its user interface. Modern ICT offers the potential, that this integration could be automated (Fig. 1). This asks for an easy access to infra red codes.

A prototype of a web based system for uploading, storing and sharing such codes has been developed as part of research work in AAL, to ease or enable automatic integration of new devices into the user interface. This prototype shows how such a system could support users and developers of ECS, AAL or smart environments in general.

IRCS also demonstrates how producers of consumer electronic could support smart environment, AAL and EC systems by making their codes available in a web based system. Applications integrating specific consumer electronic devices could download the code needed for an automatic integration of the new device.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 488 – 491, 2006.

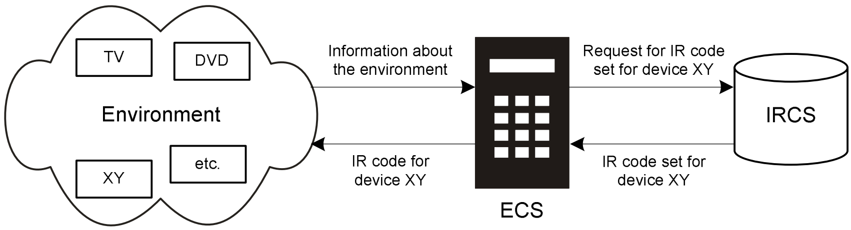


Fig. 1. Functional scheme of IRCS in a context aware system: The ECS gets information from the environment about the exiting devices. The ECS sends a request (e.g. code set for device XY) to the IRCS and then gets the requested code set(s) and allocate them to the commands. The user now can use the commands to control the device.

2 State of the Art

There are already several ECS available and most of them have a built in computer interface which allows the maintenance of them (Table 1). Unfortunately none of the off-the-shelf ECS currently allows the realization of a context aware system. An ECS is needed which allows the bidirectional communication with the environment and a server like IRCS.

Table 1. State of the Art: common Environmental Control Systems

Feature	FST James II [1]	GEWA Prog 3 [2]	Sicare Standard [3]	Proteor Nemo [4]	Autonom [5]
IR-Control	✓	✓	✓	✓	✓
Device based learning of IR	✓	✓	✓	✓	✓
PC Interface	✓	✓	✓	✓	✓

All ECS are IR based and most of them are capable to learn IR codes without the help of a computer. Before the ECS can be used it has to be set up: IR codes i.e. commands heave to be learned and often the menus have to be built. The learning process is often quite complicated since some IR codes (e.g. Philips RC5 [6]) use a toggle bit to recognise e.g. if a key of the remote control is pressed several times or for a longer time. This toggle bit can hamper the learning process of IR codes since it hardly can be recognised. Result of a wrong recognised toggle bit can be that the code doesn't work or it must be sent twice.

IRCS could also support this setup process: The user can simply download the required IR code sets from the IRCS during the setup process.

3 IRCS Functionalities

The IRCS has following functionalities and features:

- User accounting
- Category Management
- Device Management: adding new devices (TV, Radio, heater, …), duplicate check
- Code Upload
  - By users (after scanning)
  - By consumer electronic providers
- Code Download
- Code Management

The system is based on producing and storing a file using a XML DTD. The file includes all data needed for uploading. An administrator checks the codes and allows download.

Users can be assigned to different roles which grant different rights to the user. “Trusted Users” (e.g. consumer electronics industry) can give access to the codes by themselves, which means that the administrator has not to check the validity of the provided code. Codes can be downloaded, as individual files, containing one single function like “play”, or as a whole set of codes, e.g. for a particular device, by every registered user independent of the users role.

Codes are placed within categories and are assigned to a manufacturer and one of the manufacturer’s devices. Each of the previously named instances can be created by every user. Codes can be uploaded for a device by submitting an XML containing a set of codes for a device, or by entering the IR-Code for each of the devices functions. Codes entered by a non-privileged user must be validated by an administrator before they can be accessed for download. As long as the IR-Code has not been validated by an administrator the uploading user has the option to delete or alter the code. The system prevents users from entering duplicate functions and codes for one device.

Web-services allow code customisation for the target device, e.g. a handheld pc. On download the stored IR string is transformed into an IR-Code matching the devices configuration by calling one of the web-services. The transformation is not implemented yet.

Administrator or trusted users privileges are needed to maintain, change or delete categories, devices, types, functions or single codes. Registered users cannot be deleted form the system, they can only be deactivated.

4 User Interface

The system has been designed following the web accessibility guidelines WCAG 1.0 [7] and with a strong focus on usability. The design and layout can be adapted since CSS are used. Furthermore the user can choose among several deigns.

The system will be tested with practitioners (end users, care professionals) over the next months.

Fig. 2. Screen shot of the Start Page

5 Discussion

The IRCS system is a prototype of an application which forms an important part for adaptable, location based smart environments, ECS or AAL systems. IRCS should act as a repository for IR codes of consumer electronic devices which could be downloaded to interface consumer electronic devices using smart environment, AAL or ECS systems which are connected to the internet. Getting an easy or automatic access to these codes is an important contribution to more user friendly and adaptable smart environments.

The IRCS system is used to start some first negotiations with developers of ECS and AAL systems and consumer electronic devices. The system will be part of further research and development work in the field of (Smart) Environmental Control Systems (especially SmartX [8]) and AAL.

References

1. Fondation Suisse pour les Téléthèse - Stiftung für elektronische Hilfsmittel: James II, http://www.fst.ch/FST2/CATALOGUE/fiches_produits_details/James2_D.pdf
2. GEWA AB: Prog3, `http://www.gewa.se/english/english.html`
3. Dr. Hein GmbH: SiCare Standard, http://www.sicare.de/sicare_standard.html
4. Proteor Group: Nemo +)), http://www.proteor.fr/proteor/uk/ht/index_cata.htm
5. Forschungsgruppe Rehabilitationstechnik am Institut für Industrielle Elektronik und Materialwissenschaften der Technischen Universität Wien: Autonom Überblick / Einführung, http://www.fortec.tuwien.ac.at/reha.d/projects/autonom/einf/index.html
6. Seerden, P.: AN10210 - Using the Philips 87LPC76x microcontroller as a remote control transmitter, Philips Semiconductors, May 2003
7. Chisholm, W., Vanderheiden, G., Jacobs, I.: Web Content Accessibility Guidelines 1.0, http://www.w3.org/TR/WCAG10/
8. Nussbaum, G., Miesenberger, K.: SmartX open source, in: Pruski, A., Knops, H. (eds.): Assisitve Technology: From Virtuality to Reality, Proceedings AAATE 2005, IOS Press, Amsterdam 2005

The DAT Project: A Smart Home Environment for People with Disabilities

Renzo Andrich[1], Valerio Gower[1], Antonio Caracciolo[2], Giovanni Del Zanna[3], and Marco Di Rienzo[1]

[1] Centro di Bioingegneria Fondazione Don Carlo Gnocchi Onlus, Milano Italy
renzo.andrich@siva.it
[2] SIVA Fondazione Don Carlo Gnocchi Onlus, Milano Italy
[3] HB Group, Milano Italy

Abstract. The DAT project is a research initiative that aims at building up a smart home environment where people with disabilities can improve their abilities to cope with daily life activities by means of technologically advanced home automation solutions. The project has a threefold purpose. The smart home will be used as a physical setting, where clients with disabilities can follow individual programs aimed at improving their independence in the home environment. The smart house will also be used as a demonstration an educational laboratory where anybody interested can get knowledge of the latest advancements in the field of home automation and tele-care. Finally, the smart home will be used as research laboratory for testing and developing new clinical protocols and innovative solutions in the field of environmental control and home care. This article describes the architecture of the smart home, the design of the home automation system, and the research programs associated with the DAT project.

1 The Purposes of the Project

The DAT project is setting up a smart home environment where people with disabilities can improve their ability to cope with daily life activities. The project is a research initiative of the Bioengineering Centre of the Don Gnocchi Foundation.

The smart home is designed as part of a comprehensive service that also includes occupational therapy facilities and an assistive technology information and assessment service. Overall, this service will offer the opportunity for people with disabilities to go through rehabilitation and educational programs aimed at improving their independence in the home environment.

The smart home have been designed to meet three purposes. Within the Rehabilitation and Assistive Technology Counseling Services of the institute it will be used as a physical setting where people with disabilities learn and experience how to make the most effective use of technology for their independence at home. Occupational therapists and medical staff will help clients become aware of technological solutions they might adopt in their own home or residence to improve their independence. In the smart home, people with disabilities will be able to self-assess their performances and to practice independent living within training programs aimed at preparing their return at home.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 492 – 499, 2006.

The smart home will also be used as a demonstration and educational laboratory where caregivers, professionals and anybody interested can get knowledge of the latest technology advancements in the field of home automation, telemedicine, and telecare.

Finally, the smart home will be used as a laboratory for the research activities of the Bioengineering Centre of the Institute. Innovative devices, sensors, applications and protocols will be tried out and further developed for home care, independence and quality of life. The DAT smart home will give researchers the opportunity to assess the actual benefits of new technologies with final users. New protocols and devices for unobtrusive acquisition of bio-signals during everyday life activities will also be tested.

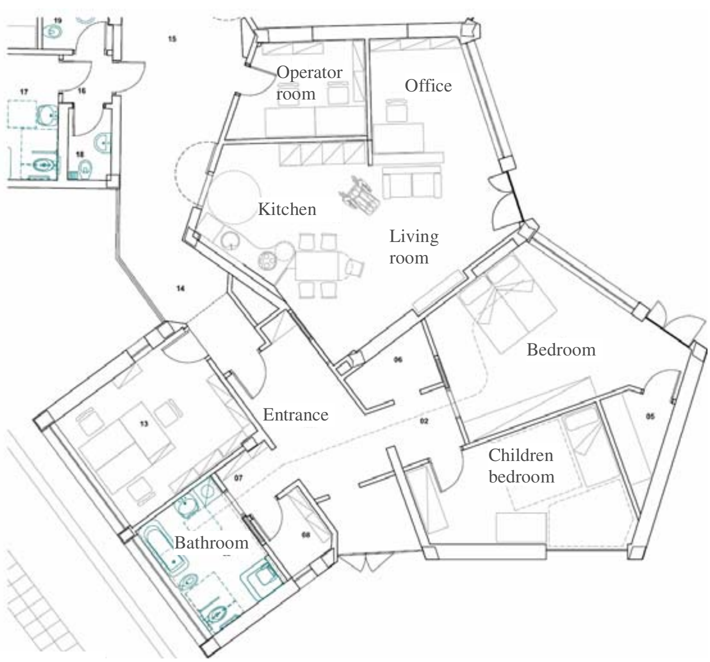


Fig. 1. Plan of the smart home

2 The Smart Home Design

When designing the home environment and choosing the home automations systems, it was very important to know about the prospective users of the smart home. The DAT project is addressed towards a broad range of clients, including people with orthopedic and neurological diseases, children as well as adults and elderly: substantially, towards any client served by the Don Gnocchi rehabilitation Institute.

A benchmarking method was used to describe this complex population. As shown in Table 1, prospective users have been divided into six different categories defined on the basis of six case studies. These include: *hemiplegia*, *paraplegia, quadriplegia*, *severe motor impairment in adult person*, *severe motor impairment in a child*, and *severe cognitive impairment*.

Table 1. Description of the six case studies

Case studies	Case description	Functional limitation
Hemiplegia	Male, 78 yeas old, stroke in the left hemisphere. Outcome: right Hemiplegia; he walks with the aid of a tripod: mild aphasia.	Autonomous walk with the aid of a tripod. No functional use of the plegic right arm. Some communication problems that increase when stressed
Paraplegia	Female, 38 years old, middle level spinal cord injury. Outcome: total paralysis of lower limbs. Upper limbs not compromised.	Unable to walk. No problem with the use of upper limbs.
Quadriplegia	Male, 58 years old, high level spinal cord injury. Outcome: total paralysis of lower limbs and upper limbs severely compromised.	Unable to walk. Unable to perform fine manipulation.
Severe motor impairment in adult	Male, 52 years old, Amyotrophic lateral sclerosis (ALS).Outcome: severe quadriplegia, tracheostomy.	Totally dependent on the caregiver. No functional use of upper and lower limbs. Impossibility of verbal expression. Minimal control of head flexion and extension.
Severe motor impairment in child	Male, 9 years old, Cerebral Palsy. Outcome: severe spastic quadriplegia.	Unable to walk. Unable to perform fine manipulation. No cognitive and communication problems
Severe cognitive impairment	Female, 65 years old, Alzheimer disease. Outcome: severe cognitive problems, light motor impairment.	Occasional motor problems but frequent loss of equilibrium. Unable to ask for help and unaware of dangerous situation. She is frightened by the presence of strangers in the home.

2.1 The Architecture of the Smart Home

The Smart Home is composed of 7 rooms: an entrance, a kitchen, a living room, an office, a bathroom, and two bedrooms, one for adults and one for children (Figure 1). The overall surface is approximately 130 squared meters. Outside the apartment, a control room allows the professional team to customize the home automation configuration and to process / analyze all bio-signals collected from clients for research purposes. Outside the apartment, there is a terrace / garden where clients will have the possibility to carry out gardening activities and test their wheelchairs on different terrains.

Of course, the house has been designed with no architectural barriers; special attention has been paid to the kitchen and the bathroom design so as to have them fully usable by people in a wheelchair (Figure 2). Particular solutions for facilitating personal mobility of people with motor disabilities inside the house have been adopted, such as sliding doors and a ceiling mounted hoist.

Fig. 2. 3D plan of the kitchen

2.2 The Automation System

The automation systems of the smart home have been chosen on the basis of the following principles. As the project has therapeutic and not only research purposes, a *commercial certified infrastructure* has been chosen to make it sure that dependability and safety requirements are met. Furthermore, the automation system had to fit a flexibility requirement to make it possible to easily change the configuration of the smart home depending on the individual case. Another important requirement was the *interoperability* i.e. the possibility to integrate products from different manufacturers into the same network. The system had also to be *easy to use* even by people that are not familiar with computers and technologies in general. A *hidden technology* concept was applied to make the house resemble a normal flat and not a research laboratory.

The chosen home automation system is based on an *EIB/konnex*™ communication protocol over a twisted pair cable. This standard indeed demonstrated, in previous projects where it has been used, a good reliability and flexibility [1]. EIB/Konnex™ is an open standard that allows interoperability, and is based on a distributed intelligence concept that makes it possible to fully work even without a computer. This last aspect is very important to fit the *reliability* requirement.

When designing the home automation system we decided to divide the automation functions into two level: the *first level automations* imply a direct relation between the input and the output (e.g. opening a door using just a remote controller); the *second level automations* involve some complex logical relation between the input and the output (e.g. automatic closure of a water tap when it is left open too long). The latter are the functions that actually make the home "*smart*" in some way.

The devices and automations installed in the smart home can also be divided on the basis of their goals. The first category includes automations primarily aimed at *improving the user's independence* in the home environment: for instance, the automation of doors, windows or lights. The second category comprises devices and automations primarily aimed at *improving safety and security* for the user: for instance, flood and fire alarms, burglary alarms or the possibility to automatically turn off dangerous appliances. Another important category is made up by the devices primarily aimed at *reducing the workload of caregivers*, like assistive devices for lifting and transferring. Finally, there are devices and automations related to telecare for *health monitoring*, such as for instance wearable sensors for unobtrusive bio-signal monitoring. Of course these categories are not rigid but rather fuzzy; belonging to one or another category depends not only on the intrinsic features of the device but also on the occasional configurations adopted in relation to a specific user's need.

3 Central Control and Distributed Intelligence

As previously stated, the chosen automation system is based on a distributed intelligence concept, in other words it can run even if the central control unit is turned off. This allows for a higher reliability, as the system is not dependent on a computer, and the possible failure of a single component won't cause all system to crash.

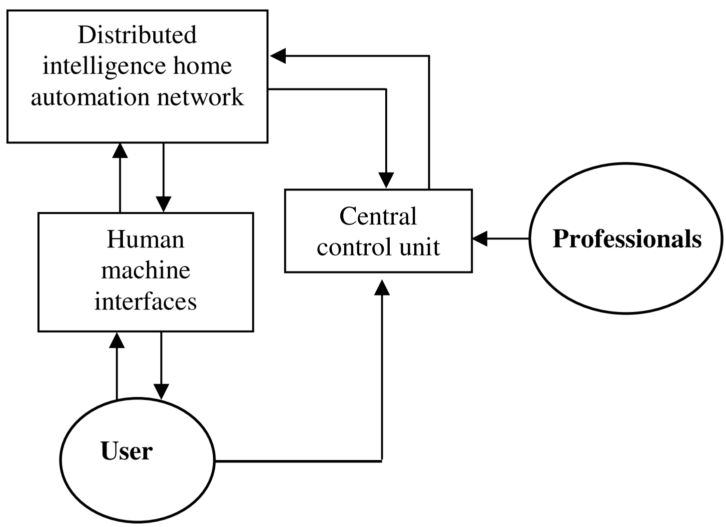


Fig. 3. The home Automation system

On the other hand, relying only on distributed intelligence would lack flexibility, which would be critical in a project that has to cope with a wide variety of prospective clients and includes research goals. Therefore the solution adopted for the DAT project is somewhere in the middle between a centralized and a distributed approach. In brief, it can be defined as a distributed intelligence system interfaced with a computer-based central supervisor (Figure 3).

The basic functions of the system (i.e. what we previously called first level automation functions) are implemented through *distributed intelligence* while more complex functions (i.e. second level automation functions) are managed by a *central control unit*. This solution ensures the required flexibility while maintaining a high dependability level for all of the basic functions.

4 The Research Program

As soon the smart home will be operational, it will undergo an experimental period for approximately one year. During this period a validation group of 30 people will be chosen among all clients, in order to try out new protocols and methods to increase the users' ability to cope in the home environment. During this period, the functionality, usability, and effectiveness of the technological solutions installed in smart house will also be tested.

After an initial assessment by the medical and the occupational therapy team, every member of the validation group will be given the opportunity to visit the smart home and have a demonstration of the automation functions installed. In this stage, clients will improve their knowledge of the possible solution to their problems in daily life activities. This knowledge improvement is an essential step for the empowerment of the user, i.e. for the process of personal growth that leads the person with disabilities to greater autonomy [2].

Every client will then be asked to formulate his/her specific problems and needs in the field of daily life activities. To facilitate this screening process, instruments such as those defined by the EUSTAT handbooks will be administered [2]. This phase will lead the user to formulate, with the help of the therapists' team, a series of objectives he/she wants (and feels prepared) to reach.

The team will then plan a series of trial sessions that may last just half a day or even more days depending on the individual case; in each trial the smart home will be configured to let the client try out the most appropriate solutions. Based on activity reports, clinical observations and interviews, the team will eventually formulate a final report with recommendations of the most appropriate solutions for the user in his/her own apartment.

The socio-economic impact of the proposed intervention will be evaluated by means of the SCAI (Siva Cost Analysis Instrument) instrument [3]. After a certain period of time (approximately 6-10 month) a follow-up will be performed using outcome measurement instruments such as the IPPA (Individual Prioritised Problems Assessment) [4], the QUEST (Quebec User Evaluation of Satisfaction with Assistive Technology) [5] and the PIADS (Psychosocial Impact of Assistive Devices Scale) [6], so as to evaluate the effectiveness and the usefulness of the proposed and/or adopted solutions.

The smart house will also be used by the researchers of the Bioengineering Centre for projects in the field of biomedical technologies and telemedicine. The smart home will serve as ideal environment for the integration of unobtrusive health monitoring devices.

Basically, two different approaches will be followed for the experiments in this field. The first approach consists of positioning the electrodes or sensors throughout the environment, namely on objects or pieces of furniture that are likely to be touched by the subject during his/her activity at home. When the contact between these "environmental sensors" and the subject occurs the signal can be recorded. The second approach, firstly proposed in 1996 at the Massachusetts Institute of

Technology, is based on the idea to include sensors and electrical connections into a garment, possibly using textile technology [7] . When the cloth is worn, the wearable sensors come into contact with the subject's body and biological signals can be detected and transmitted via a wireless connection to the smart home data bus without requiring any further instrumentation.

As part of the research program of the DAT project it is planned to use, as a tool for the unobtrusive recording of biological data, a new sensorized garment, the *MagIC* (MAGlietta Interattiva Computerizzata) developed by the Bioengineering Centre and patented in 2003 (figure 4). In brief, MagIC is a system composed of a sensorized vest and an electronic board. The vest, washable, is made of cotton and lycra™ and includes two woven electrodes, made by conductive fibers, positioned at the thorax level. The contact between the textile electrodes and the thorax is guaranteed by the elastic property of the garment. Through integrated conductive fibers, the ECG signal feeds an electronic board (having the size of a small cell phone) placed on the vest through a Velcro™ strip, which can transmits the signal through a wireless connection to a remote computer. The system simultaneously detects also the respiratory frequency and the subject's movement through a 3 axis accelerometer. MagIC has been successfully tested in laboratory and under incremental exercise [9].

Fig. 4. The sensorized t-shirt and the electronic board. The dimensions of the electronic board are compared to those of a cell phone.

Finally, the house will be used to test and further develop innovative environmental control devices based on the brain computer interface technology. The main idea of this technology is to exploit the signal produced by the motor cortex when we imagine a movement of a part of the body to control electrical devices.

References

1. Lauruska V, Serafinavicius P: Adaptative Smart Home System for Disabled. Assistive technology – Shaping the future. 579-583. IOS Press, Amsterdam 2003.
2. EUSTAT Consortium: Go for it! A handbook for users of Assistive Technology. European Commission, Milano 1999
3. Andrich R: The scai instrument: measuring costs of individual assistive technology programs. Technology and Disability 14: 95-99. Ios Press, Amsterdam 2002

4. Wessels R, Persson J, Lorentsen O, Andrich R, Ferrario M, Oortwijn W, Van Beekum T, Brodin H, de Witte L: Ippa: individual prioritised problem assessment. Technology and Disability 14: 141-145. IOS Press, Amsterdam 2002
5. Demers, L., Weiss-Lambrou, R., and Ska, B. Development of the Quebec user evaluation of satisfaction with assistive technology (QUEST). Assistive Technology, 8, (1996), 3-13.
6. Day, H., Jutai, J. W. Measuring the psychosocial impact of assistive devices: The PIADS. Canadian Journal of Rehabilitation, 9, (1996), 159-168.
7. Post ER, Orth M. Smart Fabric, or “wearable clothing”. Proceedings of International Symposium on Wearable Computers. Cambridge, Massachusetts, 1997, 167-168.
8. Castiglioni P, Merati G, Di Rienzo M: Identification of steady states and quantification of transition periods from beat-to-beat cardiovascular time series: application to incremental exercise test. Proceedings of Computers in Cardiology, IEEE Press, 2004

Selection by Visualization of Topological Layouts for Adapted Living Area Design

Arnaud Puret, Sébastien Aupetit, Pierre Gaucher,
Nicolas Monmarché, and Mohamed Slimane

Université François-Rabelais de Tours, Laboratoire d'Informatique,
Polytech'Tours, Département Informatique,
64 avenue Jean Portalis, 37200 TOURS, France
{arnaud.puret, sebastien.aupetit, pierre.gaucher, nicolas.monmarche, mohamed.slimane}@univ-tours.fr

Abstract. To design adapted houses many things must be considered: standards and recommendations related to the field of disabilities, capacities and incapacities of the persons, and wishes of the future resident. Considering all those constraints makes difficult to build such adapted living areas. Automatic layout generation simplifies the design task and decreases both design costs and study times. However, it produces a lots of layouts that we need to present to the user. In this work, we proposed clustering and visualization methods to help both the designer and the demander to choose between layouts.

1 Introduction

This study is part of the HM2PH project (Habitat Modulaire et Mobile pour Personnes Handicapées - modular and mobile house for disabled persons) [1] developed in the Laboratoire d'Informatique de l'Université François-Rabelais de Tours. The HM2PH goal is to specify a mobile and adapted living area for a disabled person. One objective of the specification is to design a movable habitat in order to allow, for example, to move the house from a medical structure to a family environment. The second objective is to built a modular house. To allow both objectives to be realized, pre-built blocks are used (walls, partitions, roof...). The main advantage of design with pre-built blocks is the cost reduction.

Softwares are used in the HM2PH project [2] to simplify adapted living area design but also to reduce both costs and design times. Other studies use softwares for the same aim [3]. The third objective is to simplify the design task such that a person that is not designer can make the study. With our approach, the following steps are necessary to design an adapted habitat:

1. study of disabilities of the future resident [4],
2. study of wishes of the future resident wishes,
3. automatic layout generation (see section 2),
4. choice of one or many layouts (see sections 3.2 and 3.4),
5. the in-house installation,
6. visit of a virtual house.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 500–507, 2006.

In this work, we will proceed in two steps. First, we present how automatic layout generation can be made. We follow by introducing how topological layouts can be visualized to help the user to choose between numerous layouts.

2 Automatic Layout Generation for Adapted Living Areas

2.1 The Constraints

The design of adapted living areas for disabled persons requires different kind of constraints. The pre-build modules add strong constraints on the habitat. The modules are very simple (rectangular shape and few different size) to reduce costs and to simplify the conception. The relative positioning of rooms is isothetic (parallel or orthogonal to external walls). Those assumptions simplify the design of an algorithm for automatically generating layouts by reducing the search space size. A minimum number of rooms must be present: bedroom, kitchen, bathroom... Moreover, some rooms are made mandatory by the Smart Home concept [5]: technical rooms for home automation, external network connections... Building costs and exploitation costs must be traduced in constraints for the layout generation. Constraints related to the future resident are of two kinds: deduced from its disabilities or fixed *a priori* (corridors width, rotation areas...) or written formally and related to the resident wishes.

2.2 Automatic Layout Generation

When the house is manually designed, too many constraints need to be handled. However, softwares allow to solve such problems and even more complex ones. Purely mathematical approach (system of equations) is difficult to use to model our constrain problem. Various methods can be used to solve the problem: genetic algorithms [6], local propagation methods [7], rule systems [8]... In our work, we decided to consider the problem as a Constraint Satisfaction Problem to solve it efficiently.

Depending on the precision of the constraint description, too many possibilities can exist and can lead to very large computation times. To reduce it, B. Medjdoub et B. Yannou [9,10] propose to split the exploration in two steps:

- a tolopogical search step: we consider only the relative positioning of rooms for which an exhaustive search is practicable,
- a geometrical optimization step: free parameters of a reduced set of layouts are optimized according to one or more criterions.

This approach has been chosen for this study. In the following, we will be particularly interested by topological layouts.

The topological layouts are not only interesting for their mathematic properties, but also for their graphical representation because they are like layout drafts. Layout drafts are easily unserstandable and there is not as much topological layouts as possible layouts. The computation time is greatly reduced since only few layouts are fully optimized (expensive operation).

2.3 Topological Layout Representation

To represent a topological layout, we chose a value for each free parameter in its associated set. Two solutions are envisaged:

– to use the middle value of the set: in this case, a topological layout has a representation very similar to a real one,
– to maximize the areas of rooms: the layout aspect is less realistic (near a draft) but the user has a better view of possible configurations for the layout (see Figure 1 (a)).

In the following, we consider the second solution: the area of rooms are maximized. Topological layouts can be numerous and can present similarities. To simplify the choice of one or more topological layouts, we propose to adapt a visualization technique.

3 Visualization of Set of Topological Layouts for Decision

3.1 Visualisation for the Decision Task

According to the precision of the house specification, the topological layout generation can produce more solutions than the user can handle. In this work, our aim is to visualize a set of topological layouts to simplify their selection. To do that, we propose to consider visualization techniques. The visualization of data greatly depends on the kind of available data. If they are vector data, lots of techniques are available [11] but if they are not, it becomes more difficult. One way to bypass the problem is to consider pairwise dissimilarity between data. With dissimilarities, more general methods are available such as multidimensional scaling [12] or pseudo-euclidean scatterplot matrix [13]. However, they are not very adapted to the decision task because they present lots of data to the user. The way to present less data to the user, or at least to organize them, is to organize data in a tree. We choose to use trees and more precisely Hierarchical Ascendant Classification (HAC) [14]. In order to efficiently use the HAC method, we need to define a dissimilarity between topological layouts.

3.2 Dissimilarity Between Topological Layouts

The Criterion. To cluster topological layouts, we need a criterion. We propose to distinguish the topological layouts through the relative positioning of rooms. When comparing layouts we do not compare only position of one room into the layout but groups of rooms. For example, all bedrooms can form a group. In the following, we consider the corridor group as an arbitrarily chosen comparison criterion because the experiments show that such criterion effectively segregate the layouts. Notice that this choice does not reduce the generality of the proposed method.

An Euclidean Distance Measure to Compare Topological Layouts. To build a measure between topological layouts we transform them into numeric values. To do that, we proceed as follow. First, it is important to note that all topological layouts have same width and same height. Let P be a topological layout. We discretise the layout in $H \cdot L$ blocks of the same size ($H \in \mathbb{N}^*$ and $L \in \mathbb{N}^*$ depend of the wanted precision of the measure). We associate an element e to each block. Let Ω be the set of the $H \cdot L$ elements associated to the blocks. Let $S(P)$ be the surface of the interesting rooms for the layout P. Let $E(S(P)) \subset \Omega$ be the set of elements of Ω associated to the surface $S(P)$.

To compare two topological layouts P_1 and P_2, we propose to consider the area resulting of the union of the surfaces minus the intersection of the surfaces given by: $(S(P_1) \cup S(P_2)) - (S(P_1) \cap S(P_2))$. This method is based on the fact that more two groups of rooms are similar (they stand in the same layout place), more the common surface is large and consequently more the uncommon area is small.

Let $|X|$ be the cardinal of the set X. From an ensemblist point of view, to compute $(S(P_1) \cup S(P_2)) - (S(P_1) \cap S(P_2))$ in the surface space is equivalent to compute $|E(S(P_1)) - E(S(P_2))| + |E(S(P_2)) - E(S(P_1))|$ in the Ω space. Comparing two layouts P_1 and P_2 using the uncommon area or comparing them using the root square of this uncommon area are two equivalent actions. In order to extend this comparison notion, we define, for a layout P and an element $e \in \Omega$:

$$f_S(P, e) = \begin{cases} 0 \text{ if } e \notin E(S(P)) \\ 1 \text{ if } e \in E(S(P)) \end{cases} \tag{1}$$

We compare the layout P_1 and P_2 with the euclidean distance calculated on the set Ω of discretized blocks e:

$$d_S(P_1, P_2) = \sqrt{\sum_{e \in \Omega} (f_S(P_1, e) - f_S(P_2, e))^2} \tag{2}$$

We show that: $d_S(P_1, P_2) = \sqrt{|E(S(P_1)) - E(S(P_2))| + |E(S(P_2)) - E(S(P_1))|}$
We generalized the function f_s with the consideration that its values are taken in the range $[0;1]$ (previously in the set $\{0;1\}$). Let $\mathcal{P} = \{P_0, \ldots P_{n-1}\}$ be the set of topological layouts then we have: $f_S : \mathcal{P} \times \Omega \rightarrow [0;1]$

This generalization makes it possible to model a fuzzy membership of an element in the surface to be compared. In this case, d_s is an euclidean distance. Then, we can consider the vector $f_S(P) = (f_S(P, e))_{e \in \Omega}$ as a representative of the surface to be compared and we can use those representatives in an euclidean space $\mathbb{R}^{|\Omega|}$. For real topological layouts, f_S takes its values in the set $\{0;1\}$ to indicate the membership or not of a block to the surface to be compared. In the following, we compute vectors f_S which have no topological layout associated and for which the vector coordinates take their values in the range $[0;1]$. Each value measures the degree of membership of a block to the surface to be compared (when a block is certainly not a member of the surface to be compared, the degree is null and when a block is certainly a member of the surface to be compared, the degree is 1). We finally obtain a numerical value measuring the difference between two layouts.

3.3 Hierarchical Ascendant Classification (HAC)

In preceding sections, we defined a way to embed topological layouts into euclidean space considering vectors. With such embedding, the utilization of the HAC [14] is simplified.

A HAC is a method to built a binary tree from data. The method proceed by grouping at each step two elements, two groups or an element with a group depending on some kind of similarity measure. When grouping, many strategy can be used such as minimal distance step or maximal distance step. An another efficient strategy is the Ward grouping criterion [15] that makes groups in order to minimize intra-group inertia. In our work, we consider the Ward criterion because it has an interesting property: it tends to build balanced trees.

When applying the HAC to the case of the topological layout classification, we must take care of some particularities of the problem. Let us take an example of particularities. We consider P_1 and P_2 two different topological layouts. Due to the embedding built for the criterion, we can have $f_S(P_1) = f_S(P_2)$. Consequently, according to the criterion, P_1 is equivalent to P_2. To reduce the bias introduced by such layouts on the HAC, we initially group those topological layouts into classes. We can note that the hierarchy may not be a binary one.

3.4 Visualization of the HAC

To visualize the HAC, we consider the NicheWorks techniques [16] and the node representatives described below.

NicheWorks Visualization. The NicheWorks technique [16] allows to visualize tree utilizing all screen space. Each node is positioned onto circles whose radius depend of the length of the path between the root and the node. The positioning algorithm uses an angular division strategy and walks through the tree from the root to the leafs. For the root node, the algorithm allocate 360° and place the node in the center of the representation. For each first-level nodes, the angle associated is proportional to the number of leafs accessible from the node in such a manner that the sum of angles is 360°. The circle is divided in arcs defined by those angles and each node is placed in the middle of the associated arc. The process continues descending to the leafs but the total angle considered is not 360° any more but the angle associated to the parent node. Figure 2 shows an application of this algorithm. This representation allows an overall view of layouts according to the criterion and highlights the quality of the construction (best constructions are achieved for balanced trees).

Node Representation. To enhance the tree visualization, we compute a representative for each node as the mean of their child nodes. The representatives are built recursively from the leafs. Let $g(n)$ be the set of the child nodes of the node n and L be the set of all leafs of the tree. If we note $c_S(n)$ the representative for the node n, we have:

$$c_S(n) = \begin{cases} n & \text{if } n \in L \\ \frac{1}{|g(n)|} \sum\limits_{x \in g(n)} c_S(x) & \text{otherwise} \end{cases} \quad (3)$$

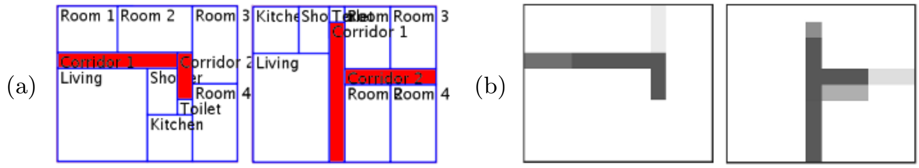


Fig. 1. (a) Drafts of two topological layouts (b) Representatives of two layout groups

With this approach, the representatives show the relative implication of each sub-tree independently of the amount of node in each branch. By construction, the average implication of a block in a node is between 0 and 1. Consequently, this node representative can be transformed into an image revealing its implications using color nuances (see Figure 1 (b)). The Figure 3 shows the representation obtained on a set of topological layouts where the criterion is based on the surface of corridors. The root node has the fuzziest representative, it gives the global repartition of the values for the criterion. The user selects one of the children, according to his idea for this criterion, to get more realistic layout. A real layout is found at the end of the tree, after few selection steps (one per tree level).

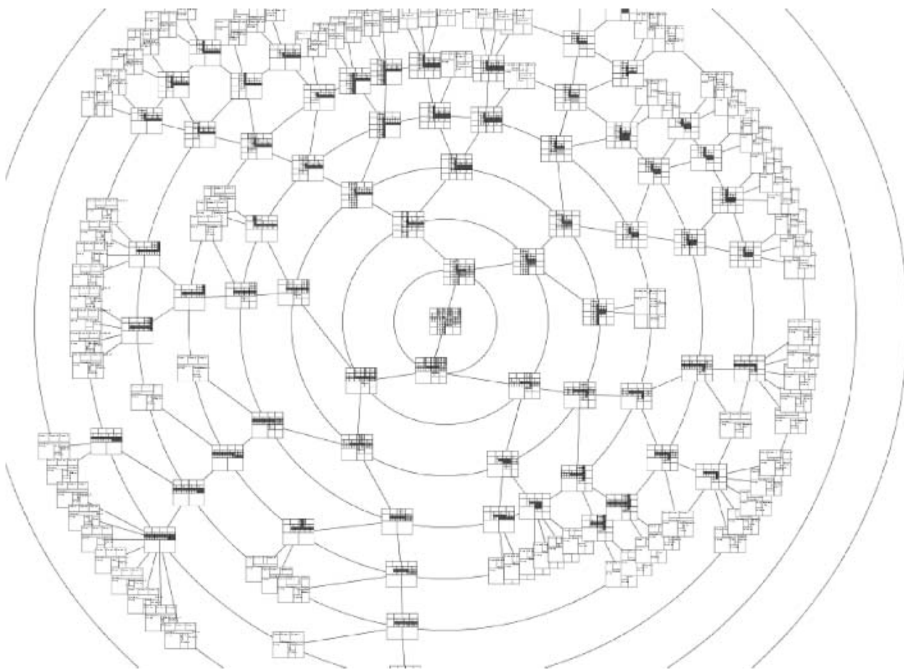


Fig. 2. Visualization of topological layouts where the criterion is based on the surface of corridors (overall view)

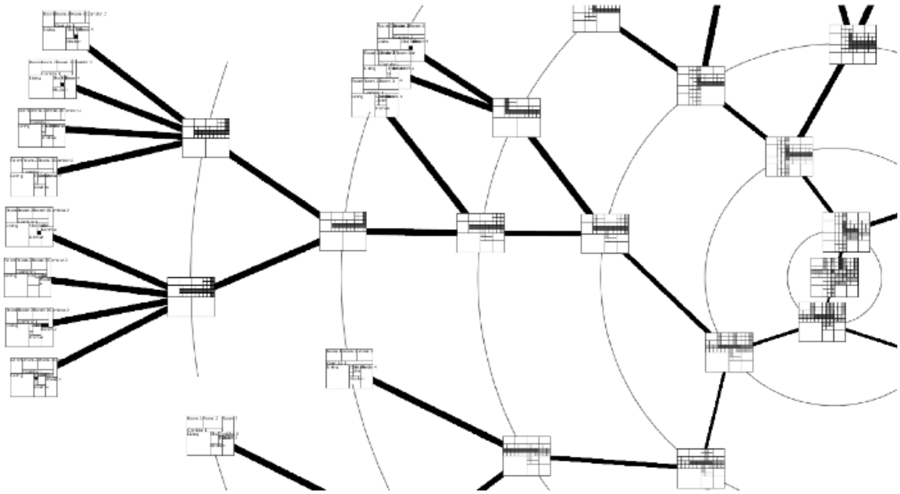


Fig. 3. Visualization of topological layouts where the criterion is based on the surface of corridors (zoomed view)

4 Conclusion

As we have quoted, the automatic layout generation process can be divided in two steps. The first one consists in generating topological layouts. The second one consists in geometrical optimisation of a reducted set of topological layouts. In order to select this reduced set of layouts, we propose to use visualization techniques. In this work, we show how topological layouts can be embedded in an euclidean space according to some criterion and how hierarchical ascendant classification can be used to build a hierarchy of topological layouts. Finally, we adapt the NicheWorks visualization method to this particular problem in order to help the user to choose one or more layouts.

Several perspectives are considered. First, we plan to widen the use of the euclidean embedding in order to manage multiple criterions simultaneously (many groups of rooms such as bedrooms and kitchen), and let the user select them according to his needs. Second, to help the user to exploit the hierarchy, we plan to add colors on edges of the tree in order to show the euclidean distance between a node and its children.

References

1. J. Leloup, *Le projet HM2PH, Habitat Modulaire et Mobile pour Personnes Handicapes, Spcifications d'un espace de vie adapt pour personne en dficit d'autonomie*, thse de doctorat, Universit Franois-Rabelais de Tours, Tours, France, 2004.
2. J. Leloup, P. Gaucher, J. Garcia and J. Siffert *The HMPH Project: Software for the Design of an Adapted Living Area*, 7th Conference of the Association for the Advancement of Assistive Technology in Europe (AAATE), 2003.

3. J.M. Martins Ferreira, T. Amaral, D. Santos, A. Agiannidis, M. Edge, *The CUSTODIAN Tool: Simple Design of Home Automation Systems for People with Special Needs*, EIB Scientific Conference, Munique, Allemagne, 2000.
4. Leloup J., Gaucher P., Pellieux S., Siffert J., *The HM2PH project, a database to help the prescription of assistive technologies*, 9th International Conference on Computers Helping People with Special Needs (ICCHP 2004), 2004,
5. P. Gaucher, *Chap. 6, Assistance aux actions sur l'environnement, dans Assistance Technique au Handicap*, coordination A. Pruski, sous l'gide de l'IFRATH, Herms, 2003.
6. J.H. Jo, J.S. Gero, *Space layout planning using an evolutionary approach*, Artificial Intelligence in Engineering Design, Volume 11, Pages 245-262, 1998.
7. G. Kwaiter, *Etude et dveloppement d'un modeleur dclaratif 3D temps rel d'environnements virtuels, bass sur les mtaheuristiques issues de la recherche locale*, thse de doctorat, Universit Paul Sabatier de Toulouse, Toulouse, France, 1998.
8. D. Plemenos, *Contribution l'tude et au dveloppement des techniques de modlisation, gnration et visualisation de scnes*, thse de doctorat, Universit de Nantes, Nantes, France, 1991.
9. B. Medjdoub, B. Yannou, *Separating topology and geometry in space planning*, Computer-Aided Design, Volume 32, Issue 1, Pages 39-61, 2000.
10. B. Medjdoub, B. Yannou, *Dynamic space ordering at a topological level in space planning*, Artificial Intelligence in Engineering, Volume 15, Issue 1, Pages 47-60, 2001.
11. S. K. Card, J. D. Mackinlay, B. Shneiderman, *Readings in information visualisation: using vision to think*, Series in interactive technologies, Morgan Kaufmann Publishers, 1999.
12. Borg, I. and Groenen, P., *Modern multidimensional scaling: theory and applications*, Springer, Springer series in statictics, 1997,
13. S. Aupetit, *Contributions aux modles de Markov cachs : mtaheuristiques d'apprentissage, nouveaux modles et visualisation de dissimilarit*, thse de doctorat, Universit Franois-Rabelais de Tours, Tours, France, 2005.
14. J.B. MacQueen, *Some methods for the classification and analysis of multivariate observations*, Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1, Pages 281-297, 1967.
15. J. H. Ward. *Hierarchical grouping to optimize an objective function.*, Journal of The American Statistical Association, Volume 58, Pages 236-245, 1963,
16. G. J. Wills, *Nicheworks - interactive visualization of very large graphs*, Journal of Computational and Graphical Statistics, Volume 8, Number 2, Pages 190-212, 1999.

Wireless Communicator for Patients in Intensive Care

Cecília Sik Lányi and Viktor Magyar

Virtual Environment and Imaging Technologies Laboratory
Department of Image Processing and Neurocomputing, Pannon University
H-8200 Veszprém, Egyetem u. 10., Hungary
lanyi@almos.vein.hu, magyarviktor@gmail.com

Abstract. Intensive care is often a stressful experience for patients and presents unique challenges for the nursing staff. Under conditions where patients are unable to speak (e.g. mechanical breathing, intubation) it becomes difficult for nurses to be aware of patients needs. We have developed two multimedia software programs that are designed to assist patient/nurse communication in these situations.

1 Introduction

Augmentative and Alternative Communication (AAC) refers to ways (other than speech) that are used to send a message from one person to another [19]. AAC strategies assist people with severe communication disabilities to participate more fully in their social roles including interpersonal interaction, learning, education, community activities, employment, volunteerism, care management, and so on [20]. In numerous of the studies, Light-tech system (e.g. alphabet boards, printed copies of overlays) were compared to voice output communication aids [3-5]. In seven of the nine studies no effect for light-tech versus high-tech AAC system on listeners' overall attitudes was found. The evidence suggests that an AAC system is probably not a single controlling factor in attitudes toward individuals who use AAC [14]. Of the nine studies, in only one was it reported that a high-tech option elicited more positive attitudes than light-tech system on all subscales of the attitude instrument [5,14].

But the situation is not so simple for a patient, who is on the Intensive Care (IC), because he/she was able to speak earlier, and now he/she is suddenly speechless. IC can be a stressful experience for patients. Many short and long term aspects of patient well being can be adversely affected following a negative IC experience. The difficulty for patients in communicating their needs to health care staff can negatively impact their IC experience. Most medical and nursing staff assume that communication problems only affect the patient during the acute treatment period but there is evidence that, even after discharge from hospital, IC-related sequelae can continue to affect the psychological well being of many patients [6,7,13,17,18].

Patients in intensive care are often unable to communicate with their environment. There are several reasons for this. Many patients require mechanical breathing assistance during intubation and this limits the patients ability to verbally communicate. In

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 508 – 515, 2006.

such situations it is a very difficult task for even the most experienced and skilled critical care nurse to communicate with the patients and truly understand their needs. Several alternative communication methods exist including alphabet boards, pen and paper, and word mouthing These methods are usually time-consuming, inefficient and frustrating for both patients and nurses. It may take a long time for the nurse to understand what the patient requests and when the nurse finally comprehends the patients needs, she or he has to go to bring the required equipment (for example: medicine, drug, bed-pan) [1,2,11,12].

2 Objectives

We have developed two software programs over the last two years to help improve patient/nursing staff communication in the intensive care ward. Both software programs have been tested out in the Department for Heart Surgery at the Hospital of Zalaegerszeg [9,10]. But in this age of dramatic advancements in wireless networks and pocket PC technology, we decided to evolve this software with a wireless component. The main goals for the software to minimise the required time to tend the patients' needs and enhance the level of patient comfort and care. Other groups have attempted to support nurse patient communication via the application of computer aided approaches, but these ones are not support mobility [8]. So we put a wireless internet connection into the focus of the current software programs. It send e-mails to a PDA in the pocket of the ward nurse, and signals the nurse that the patient needs something. The nurse can then determine the patients needs and more efficiently arrive at the bed with the required equipment.

To accomplish these objectives, two software packages have been created: The Wireless Communicator and The Virtual Mail Editor system. The Wireless Communicator was developed for patients who are transiently incapable of speech, to enable them to communicate with nursing healthcare staff and inform them on the state of their health. It was designed to help patients to communicate with their environment who are conscious, but due to some reason, (e.g. they are on a respiratory engine) are unable to speak. Using a single key, similarly to a nurse-calling bell, these patients can actuate the system. The Virtual Mail Editor Software (VME) is designed to enable patients in intensive care to communicate, with their relatives or send an e-mail using a virtual keyboard. Patients can see the virtual soft-keyboard on an LCD display. They have to manipulate only one single button on the virtual soft-keyboard to print the desired character. The system features an adjustable speed moving frame that scrolls through a set of characters and allows the patient to press a button to choose the character currently within the frame. This button acuates communication to one of several internet-connected desktop computers remotely located outside of the intensive care ward.

The software has been created in English, German and Hungarian. After writing a message or an e-mail text this message can be sent to another workstation or even in the form of an e-mail to an external receiver. Since intensive care wards are typically overcrowded with various types of equipment [15], special care was taken to develop a system that had low space requirements.

3 Methods

3.1 Description of the Software of the Wireless Communicator

The Wireless Communicator system enables the patients to select words from a large menu-window in order to communicate their wishes with the doctors and the nurses. The selected message is sent by way of an e-mail message. The message arrives to a PDA carried by the ward nurse. The messages are also stored in a database on the patients' computer and this supports documentation of the state of health of the patients [10].

Nurses and doctors in different hospitals were consulted on the creation of a database that contained the most relevant words needed for communication of fundamental patient needs. Standard communication methods as practised at the Heart Surgery department of Zalaegerszeg Hospital were also gathered from the suggestions of nurses and convalescing patients. The discussions with this latter group was very important, as these patients could provide a wealth of advise based on their own experiences in intensive care.

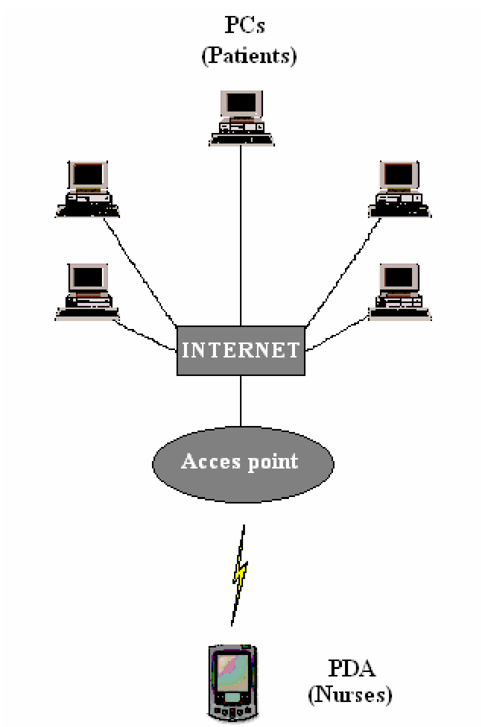


Fig. 1. Network structure

The PCs located at bedside allow patients to send e-mails via the internet (Figure 1.) to a PDA which is always in the possession of the responsible ward staff. This supports a more immediate communication of the ongoing needs of the patient to the staff.

The most difficult problem we encountered during the realisation of this system was how to harmonize the requirements of the doctors with the possibilities that such informatics offers. Since patients may at times fade in and out of consciousness, it was important to provide a clear and uncomplicated representation of the different care groups that a patient might have on the screen. In the design of this visualisation, we had to avoid typically popular spectacular elements, and instead concentrated on using the most simple and easily comprehensible realisations. It was important to select pictograms that were self-explanatory and simple (see Figure 2).

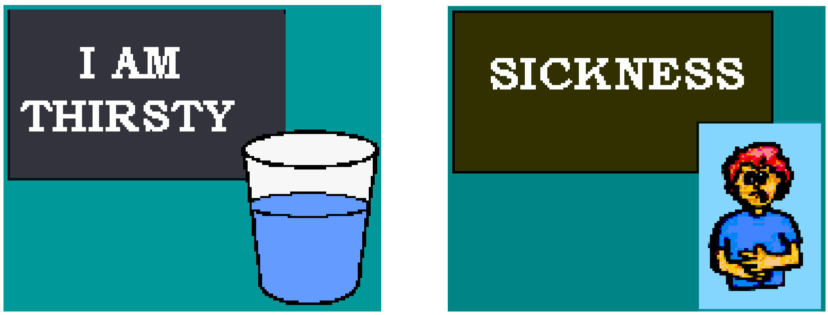


Fig. 2. Illustrations

These selected pictures were prepared to illustrate common patient needs and were positively accepted during the initial user tests. We also gathered information from the physical tools used in the wards, and tried to model them in our applications [14]. For example, one such device was the “pain level indicator”, a simple paper ruler that was used to help the patient to indicate the intensity of the pain they were experiencing. In our realisation, this approach has been simulated with a series of ten pictures, shown in a sequence which corresponds to increasing levels of pain. This enabled the patient to push a button when he or she thought that the pain shown on the figure corresponded to what he/she had felt. We chose to use such push-button devices since they were familiar to the patient (very similar to the nurse calling push-button) and were relatively easy to acuate.

A novice user can easily use the main menu of the software too. Using the push-button device, the patient can select one of the groups of requirements and is then presented with sub-groups of information that is transmitted by the software using email messaging. Object selection from the single groups is done in the following manner: in a sequence every group becomes visible in expanded form until it finally fills the entire screen until the illustrated picture becomes visible. If a group has been selected, then the different requirements within that group become visible and the patient’s selection is performed in a similar fashion. The e-mail, which contains the selected requirement, then arrives to the nurses’ e-mail address. As soon as the nurse becomes aware of the patient request, he or she can register this with a click and the requirement gets stored in the database of the actual patient. The nurse can then go to the patient already knowing what his/her request is, even if the patient is temporally unable to speak. Using such a wireless network and PDA system may ease the ward nurses’ work since they don’t have to sit in front of the server computer waiting for the patients' messages. The ward nurse has a PDA in his or her pocket and receives an immediate signal when the e-mail arrives.

The software also enables a record to be generated that documents the rehabilitation activities of the patient. At the present moment the text of the Communicator can be switched to both English and German language modes. Thus, patients who do not speak Hungarian, but understand one of the other two languages, can also communicate more effectively with the nursing staff. At the patients' bedside the PC text is seen in English or German, while the output language of the e-mail arrives to the nurses is Hungarian. As the system uses a data-base of expressions related to codes, it is a simple matter to translate this type of information into other languages as well.

3.2 Description of the Software of the VME (Virtual Mail Editor)

The VME software supports communication between the patient and the outside world. The patient can compose e-mail messages with the help of a simple button interface. The patient has only to wait till a boarder gets to the character that he or she would like to select, and then push the button. Figure 3 shows the moving boarder (seen as a blue quadrangle around one of the virtual keyboard button). In this manner, the patient can very easily compose a message. The software then sends the message as an e-mail.

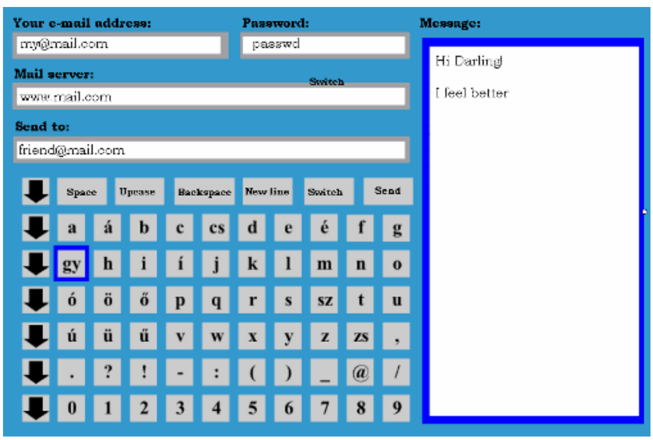


Fig. 3. Virtual Mail Editor Software

The program has been written to serve patients who speak Hungarian, German and English, but it is very easy to expand its data-base to other languages. The database is a simple text file, containing the each character, special key.

The programs are not very demanding on machine resources, as they do not use any animations or complicated digital content. Generally, the software and hardware requirements are also quite basic in order to maximize greater access to these systems (i.e., Windows 98 / 95 / NT, 200 KHz processor, 64 Mb RAM, LCD monitor, internet connection, at the nurses' side PDA, wireless internet connection, software for reading e-mails).

4 Results

4.1 Evaluation of the Wireless Communicator

We have conducted initial tests of the system at the Hospital of Zalaegerszeg, within the Heart-Surgery, Anasthesiologia and Intensive Care Departments. A questionnaire was developed to acquire feedback from the doctors and nurses at the test site. The questionnaire included items on basic demographic information (i.e. name, age, status and knowledge with computers) and on opinion on ease of use. Participants were also queried on what aspects of the system were unnecessary and what features would be desirable in future iterations of this applications.

This questionnaire was answered by 12 doctors, 19 nurses and 15 patients. The members of the last group had been completely unfamiliar to computers, while the members of the former two had used only basic documentation functions of computers and the internet, but never used programs similar to the concerned twins. Examples of user feedback comments that were received most often included:

- it can be learned within a few minutes,
- there might be problems how to accommodate the single equipment,
- it can be used in special areas,
- it has to be publicised first,
- it should still be developed further.

The questionnaire results generally indicated that it is sometime quite challenging to introduce new methods of multimedia communication in this setting. The circumstances that are characteristic in a hospital setting are uniquely different from those encountered in other application areas. Consequently our results underscored the value of collecting user centered data early on in the design process to guide iterative development with the needs of the targeted user group clearly in mind.

4.2 Evaluation of the Virtual Mail Editor Software

This software was also tested at the Heart-Surgery and Anasthesiologic Intensive Therapy Department at the Zalaegerszeg Hospital with the same test-group of doctors and nurses in similar fashion to the Wireless Communicator evaluation described above.

Examples of user feedback comments that were received most often included:

- It can be learned within a few minutes,
- the physical placement will be difficult,
- it can be used in special wards (e.g. septic wards, in wards where the patient has to stay for a long period),
- it has to be publicised,
- some development is still needed, the database of words has to be enlarged,
- it can be probably used in the future where young patients are involved, who are mentally intact,
- at present the mobile phone is sufficient.

We do not agree with this last statement since it disregards the fact that a patient who is temporarily unable to speak cannot use the mobile phone but would be capable of composing a message using our virtual soft-keyboard interface and send it as an email message. It could be good communication equipment also in cases when due to some epidemic reasons the patients can not be visited.

5 Summary

We have developed two software programs and interfaces for helping patients in intensive care departments to communicate with nursing care staff and people outside of the hospital setting. Both systems have been tested with nurses and doctors at the Heart Surgery Department and Anasthesiologic and Intensive Care Departments of the Hospital of Zalaegerszeg. The programs are capable of displaying words and expressions written in English, German and Hungarian. The software has minimal hardware requirements with the most expensive component being an LCD monitor and PDA that was selected in consideration of the space limitations that typically exist in an Intensive Care Department The twin-programs proved well in the Anasthesiologic and Intensive Care Departments of the Hospital of Zalaegerszeg, which is more than encouraging for us to introduce it in other health institutions all over Hungary and even abroad.

Acknowledgements

The authors would like to acknowledge the help of Mrs. Ildikó Müllerné Szögedi, who carried out the testing at the Heart Surgery Department and Anasthesiologic and Intensive Care Departments of the Hospital of Zalaegerszeg.

References

1. Albarran JW. (1991). A review of communication with intubated patients and those with tracheostomies within an intensive care environment. Intensive Care Nursing, 7: 179-186
2. Ashworth P. (1980) Care to communicate. London: Royal College of Nursing.
3. Beck A., Bock S., Thompson JR., Kosuwan K. (2002). Influence of communicative competence and augmentative and alternative communication technique on children's attitudes toward a peer who uses AAC. Augmentative and Alternative Communication, 18: 217-227.
4. Dada S., Alant E. (2002). A comparative study of the attitudes of teachers at special and educationally inclusive schools towards learners with little or no functional speech using communication devices. South African Journal of Education, 22(3): 223-228.
5. Gorenflo CW., Gorenflo D. (1997). Effects of synthetic speech, gender, and perceived similarity on attitudes toward the augmented communicator, Augmentative and Alternative Communication, 13: 87-91.
6. Halfsteindottir TB. (1996). Patient's experiences of communication during the respirator treatment period. Intensive and Critical Nursing, 12: 261-271
7. Hemsley B., Sigafoos J., Balandin S., Forbes R., Taylor C., Green VA., et al.(2001). Nursing the patient with severe communication impairment. Journal of Advanced Nursing, 35: 827-835.

8. Judson A., MacAulay F., Etchels M., Ricketts, IW., Waller A., Brodie J., Warden A., Alm N., Shearer A., Gordon B. (2002). Developing ICU-Talk. A computer based communication aid for patients in intensive care. Dept. of Appl. Comput., Dundee Univ.; Euromicro Conference, Informatics: Inventing the Future. IEEE Comput. Soc. 2: 370-6.
9. Kosztyán Zs., Sik Lányi C., Müllerné Szögedi I. (2002). Virtual Message – A software for the sick in the intensive care. Die Kibernetik in die Theorie und Praxis, Den 7. Internationalen TAKIS-Kongres, Nitra, Sept. 2-5. 2002, pp. 281-285.
10. Kupeczik I., Gyurka Sz. (2002). Developing communication aids "Communicator-software", Young Scientific Students Work, University of Veszprém
11. Leathart AJ. (1994). Communication and socialisation (1): an exploratory study and explanation for nurse-patient communication in an ITU. Intensive and Critical Care Nursing, 10: 20-24.
12. Lenore E., Ogle KR. (1999). Nurse-patient communication in the intensive care unit: a review of the literature. Aust Crit Care. 12(4):142-5.
13. Maddox M., Dunn SV., Pretty LE. (2001). Psychosocial recovery following ICU: experiences and influences upon discharge to the community. Intensive and Critical Care Nursing. 17: 6-15
14. McCharthy J., Light J. (2005). Attitudes toward Individuals Who Use Augmentative and Alternative Communication: Research Review. Augmentative and Alternative Communication, 21: 41-55.
15. McCrindle R., Arnold FG., Cook J., Barett D. (2002). Booy: Inclusive design of an interface a hospital system for information, communication and entertainment. The 4th International Conference on Disability, Virtual Reality and Associated Technologies, Veszprém, 18-20 Sept. 2002, pp. 139-148
16. Müllerné-Szögedi I. (2001). Media – Informatics - Communication – Care, "Art of silence", Media Informatics Communication Conference, Veszprém, 2001 Sept. 13-15, pp. 209-231.
17. Russel S. (1999). An exploratory study of patients' perceptions, memories and experiences of an intensive care unit. Journal of Advanced Nursing. 29: 783-791
18. Weiss H., Rastan V., Mullges W., Wagner RF., Toyka KV. (2002). Psychotic symptomps and emotional distress in patients with guillian-barre syndrome. European Neurology 47: 74-78
19. The American Speech – Language – Hearing Association website: `http://www.asha.org/public/speech/disorder/Augmentative-and-Alternative.htm`
20. The Augmentative and Alternative Communication Centres website: `http://acc.unl.edu/`

The Development and Application of the Assistive Devices Application System for Clients with Disabilities

Tzyh-Chyang Chang[1,2], Jiann-Der Lee[1,*], Shwu-Jiuan Wu[3,4], and Juei-Fen Huang[5]

[1] Department of Electrical Engineering, Chang Gung University, Tao-Yuan, Taiwan
259 Weng Hwa 1st Rd., Kweishan, Taoyuan 333, Taiwan
{tcc200, jdlee}@mail.cgu.edu.tw
[2] Department of Occupational Therapy, Chang Gung University, Tao-Yuan, Taiwan
259 Weng Hwa 1st Rd., Kweishan, Taoyuan 333, Taiwan
tcc200@mail.cgu.edu.tw
[3] Graduate Institute of Medical Informatics, Taipei Medical University, Taipei, Taiwan
250 Wu-Xin Street, Taipei 110, Taiwan
sjwu123@yahoo.com.tw
[4] Taipei Veterans General Hospital, Taipei, Taiwan
201, Sec. 2, Shih-Pai Rd., Taipei 112, Taiwan
sjwu123@yahoo.com.tw
[5] Department of Foreign Language, Kinmen Institute of Technology, Taiwan
1, Dasyue Rd., Jinning Township, Kinmen County 892, Taiwan
jueifen123@hotmail.com

Abstract. It is important for clients with disabilities to live independently and enhance their recovery progresses by having assistive devices in their daily lives. The purpose of this study is to develop and construct an inquiry and application system for clients with disabilities so that the officials can better manage assistive devices application information. In addition, we can check and understand how well assistive devices assist clients with disabilities in their livelihood and do an analysis based on current conditions and results through the system. These results can then be used as references for improving future clients with disabilities welfare suggestions to government officials.

1 Introduction

Assistive devices are vital vehicles for clients with physical disabilities to gain independence from daily lives and enhance their recovery progress. Clients with disabilities using assistive devices not only regain independence from daily lives but also improve their live and work performance thus greatly reduce care-taker's burden [1-3]. Due to increasing demand of assistive devices, new research and design on them and governmental subsidization to produce and repair them have been a critical issue facing the government [4]. When clients with disabilities need assistive devices, they are required to apply through district offices which in term report back to social welfare bureaus of different cities or counties. City and county social welfare bureaus' clients with disabilities assistive devices application system can only accept

* Corresponding author.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 516 – 523, 2006.

application but can not meet effectively different needs of clients with disabilities, nor can it accurately predict budgets necessary for purchasing suitable assistive devices for clients with disabilities because of time restriction. Since assistive devices are of great assistance to the clients with disabilities in improving their daily lives, many of them apply actively for the instruments accordingly. It would be to the handicapper's best interest if different social welfare bureaus could approve those applications in the shortest time limit possible. In addition, understanding the proper functions each device serves to clients with disabilities, effective management of those assistive devices, as well as precise budgeting all could contribute to future governmental welfare plan for clients with disabilities.

The concept of Decision Supportive System was proposed in the early '70, it was called "Management of Decision System." Its main characteristic is to assist users to use data and solve unstructured questions by using computer technology, thus further enhances computer application gradation. By utilizing computer technology, not only traditional electrical data processing will be carried out, it also provides assistance to users making daily decisions [5-6]. Decision Supportive System comprises of three main segments: user interface, rule database, and knowledge database. User interface provides user-friendly introduction and consultation functions. It usually will be designed differently according to various different implicational needs. Rule database holds many deduction rules and controls the whole deduction process; it will retrieve solutions from existing data in the knowledge database. Knowledge database saves existing data or experienced data gathered from human experts or it can also deduct understanding knowledge from the process of programming [5-6]. Furthermore, Decision Supportive System not only provides health-care providers with diagnostic reference but also provides general public with health consultation if used in medical related applications. It can also serve the purpose of training professional medical personnel through multi-system integration [7-8].

We, therefore, can apply the same concept and structure of Decision Supportive System to assistive devices management system. With the convenience of inquiry system, the desired data can be found with ease by searching into the existing database. By doing so, it is expected that users, administrators, or medical-related personnel will work with efficacy in handling handicappers' assistive devices applications, assistive devices management and clinical data retrieve [9-10]. On Jan. 16, 2004, Ministry of Interior revised the standard chart of assistive devices for clients with disabilities. The primary purpose of the revision is to gain better use of assistive devices resources for clients with disabilities. This research team uses assistive devices application data from Taipei city social welfare bureau from Jan. 1, 2000 to Dec. 31, 2003 to analyze and construct an inquiry and application system for clients with disabilities.

The purpose of this study not only intends to develop an inquiry and application system for clients with disabilities assistive devices but also expects to exam the suitability of Taipei city government subsidy for the instruments before and after the revision. Through the use of the system, we can better understand and analyze current improvement made by incorporating assistive devices in handicappers' lives and provide government with reference to improving clients with disabilities future social welfare.

2 Method

The sampling data is collected and analyzed from Taipei city social welfare bureau from 2000 to 2003. There were a total of 41,900 people (male 56%; female 44%) applying for assistive devices during these four year periods (see table 1).

Table 1. Comparison of assistive devices application for each year

Year	2000		2001	2002	2003
applicant	8041		10303	11635	11921
Sex	Male	57%	56%	56%	56%
	female	43%	44%	44%	44%

According to Taipei city government regulations, clients with disabilities assistive devices application submission, verification and approval need to be done directly in district offices. Every application and subsidy, according to the regulations, is categorized into low-income household and non low-income household with maximum amount of subsidy available, minimum years of possession, subsidiary populations and subsidiary measures concluded.

Clients with disabilities in Taipei city can apply for assistive devices with the assistance of district offices. District offices can apply for assistive devices requested by applying to social welfare bureau's on-line system (see figure 1).

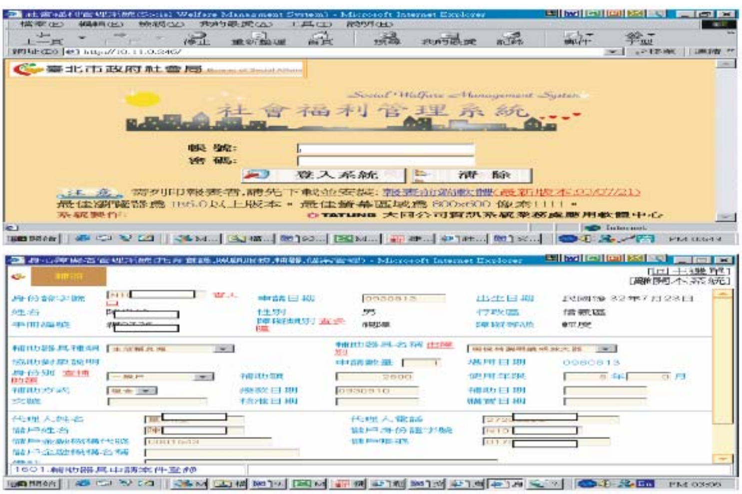


Fig. 1. Assistive devices subsidy management system

Taipei city social welfare bureau's clients with disabilities assistive devices application system is constructed under social welfare management system of Taipei city government. It can only support applications for assistive devices for clients with disabilities. It can not supply management of assistive devices, nor can it provide an easy and effective means of inquiry for frequently-demanded assistive devices and subsidy. Moreover, the application system for assistive devices is not constructed independently from the social welfare management system of Taipei city government; it is, in fact, quite difficult for administrators to adjust functions that are related to assistive devices inquiry.

It hence becomes our aim to develop an inquiry and application system for clients with disabilities assistive devices to examine whether there's any difference before and after the revision of government regulations, and the suitability of subsidy as well as to better understand and analyze the livelihood of clients with disabilities with the assistance of assistive devices.

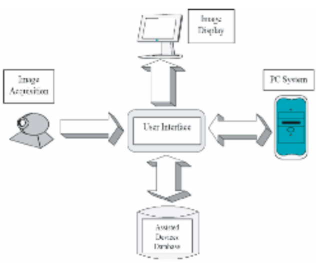


Fig. 2. The structures of this research

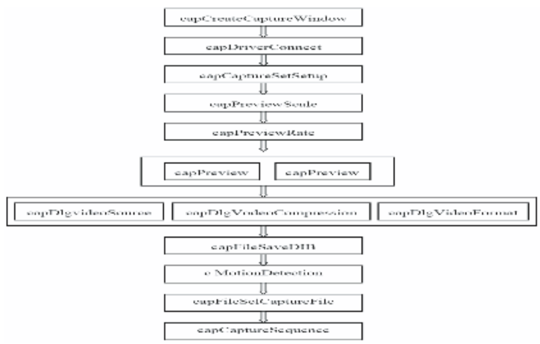


Fig. 3. VFW structure chart

The structures and consists of this study (see figure 2) are as follows:

1. Assistive devices application and inquiry system: the related programming for the system and user interface design was included in this study.
2. Programming software: Borland C++ Builder 6.0 was used in this study. This software is not only easy to operate but also is easy to integrate database. Furthermore, it is portable as well.
3. Database inquiry system: Microsoft Access 2002 was used in this study. Assistive devices information and user information search are all incorporated in the database of this system. Users can update database content according to different conditions.
4. Image acquisition interface: uses VFW (Video for Windows) SDK designed specifically for software developers by Microsoft. Figure 3 showed the VFW consists of parts of this system.

3 Results

The preliminary application and inquiry system this research team designs is principally for clients with disabilities applying for assistive devices (see figure 4). There are function bars for selection on the very top of the screen. Users can key-in any one piece of information of clients with disabilities in single folder of the database and get subsidy information for that person on assistive devices. Besides, this system also provides another function – image acquisition. That means it can provide images of assistive devices to applicants or use the web cam as a means for assistive devices usage counseling in the future.

Fig. 4. Assistive devices inquiry system designed by this research team

In clinical use of the system, users can key-in search key in the search function (including AND and OR), another window will show the assistive devices result based on the search demand. In addition, they can also search for assistive devices subsidy information or specific assistive devices for certain handicap. For example, we set our search question as “how many males are wearing Both Ears Audiphone?”, the result indicates that there are 1,878 items (see figure 5). Administrators can also change or update content of data bank information or assistive devices applicants’ information after entering passwords.

Fig. 5. System showing information after receiving inquiry information

4 Discussion

After the preliminary assistive devices application and inquiry system is finished, we analyze the target applicants from 2000 to 2003 and get the following data analysis:

According to the four-year time span statistics, the top five categories of assistive devices clients with disabilities applicants applied for are (1) Wheelchair: 6,427 applicants (15%), (2) Both Ears Audiphone: 6,248 applicants (15%), (3) Cushion Bed Mattress : 5,744 applicants (14%), (4) Raised Toilet Seat: 1,942 applicants (5%), (5) Wheelchair Cushion Seat: 1,574 applicants (4%). Of all the assistive devices applied by clients with disabilities, the Both Ears Audiphone seems to be the most popular one (with subsidy up to US$ 875 per person) (see table 2).

Table 2. Top five categories of assistive devices clients with disabilities applicants apply for

Rank	Assistive Devices
1	Wheelchair: 6,427 applicants (15%)
2	Both Ears Audiphone: 6,248 applicants (15%)
3	Cushion Bed Mattress : 5,744 applicants (14%)
4	Raised Toilet Seat: 1,942 applicants (5%)
5	Wheelchair Cushion Seat: 1,574 applicants (4%)

Table 3. Annual comparison of the top three assistive devices categories for clients with physical disabilities applicants

Year	2000	2001	2002	2003
Applicants (%)	36%	38%	38%	39%
Rank 1				
Assistive Device	Wheelchair	Wheelchair	Wheelchair	Wheelchair
Applicants	457	643	801	808
Rank 2				
Assistive Device	Cushion Bed Mattress	Cushion Bed Mattress	Cushion Bed Mattress	Cushion Bed Mattress
Applicants	315	477	618	661
Rank 3				
Assistive Device	Specially Made Tricycle	Wheelchair Cushion Seat	Raised Toilet Seat	Raised Toilet Seat
Applicants	192	203	301	308

On social welfare policy, this research team also would like to learn current conditions different handicapped applicants faced applying for and using of different assistive devices. Through the use of the system, we conclude the top three handicapped categories and their respective assistive devices needs as follows:

1) Clients with physical disabilities: Wheelchair (18%), Cushion Bed Mattress (14%), Raised Toilet Seat (6%) (see table 3);
2) Clients with multiple disabilities: Cushion Bed Mattress (19%), Wheelchair (13%), Both Ears Audiphone (6%) (see table 4);
3) Clients with hearing impaired: Both Ears Audiphone (56%), One Ear Audiphone (10%), Fax Machine (4%) (see table 5).

Table 4. Annual comparison of the top three assistive devices categories for clients with multiple disabilities applicants

Year	2000	2001	2002	2003
Applicants (%)	27%	28%	28%	27%
Rank 1				
Assistive Device	Cushion Bed Mattress	Cushion Bed Mattress	Cushion Bed Mattress	Cushion Bed Mattress
Applicants	352	547	580	672
Rank 2				
Assistive Device	Wheelchair	Wheelchair	Wheelchair	Wheelchair
Applicants	287	402	461	402
Rank 3				
Assistive Device	Both Ears Audiphone	Both Ears Audiphone	Raised Toilet Seat	Phlegm sucking machine
Applicants	142	174	193	222

Table 5. Annual comparison of the top three assistive devices categories for clients with hearing-impaired applicants

Year	2000	2001	2002	2003
Applicants (%)	27%	20%	18%	18%
Rank 1				
Assistive Device	Both Ears Audiphone	Both Ears Audiphone	Both Ears Audiphone	Both Ears Audiphone
Applicants	1172	1410	606	1539
Rank 2				
Assistive Device	One Ear Audiphone	One Ear Audiphone	One Ear Audiphone	One Ear Audiphone
Applicants	202	215	211	227
Rank 3				
Assistive Device	Fax Machine	Fax Machine	Fax Machine	Shaking type watch
Applicants	107	84	77	73

5 Conclusion

The assistive devices application and inquiry system by this research team designed is still under development. Through search result analysis utilizing this system, we get more comprehensive understanding towards current assistive devices application conditions of clients with disabilities and hope that this may in term assist the government in improving future welfare for clients with disabilities. It is expected that this system may provide more suitable services, secure assistive devices assessments, and custom-made rehabilitation training in the near future.

References

1. Ko, I.J., Cho, B.Y., Yang, S.C., Chang, W.I., & Siu, F.J.: The Project of Assistive Devices Lending. Journal of Long Term Car, 8(3) (2004) 345-354
2. Wu, W.D., Wang, H.P.: The Enhance of Assistive Technology for Clients with Disabilities. Journal of Special Education, 72 (1999) 1-9
3. Chen, J.J.: The Need and Design for the Elders' Daily Assistive Devices. The Journal of Min-Chi Technique School, 32 (2000) 163-174
4. Yen, C.C.: The Assessment of Assistive Technology for Disabled Students. Journal of Life Technique Education, 32(5) (1999) 25-32
5. Lee, Y.C., Liu, L., Chu, W.D.: Medical Diagnostic Supportive System. Journal of Taiwanese Medicine, 1(6) (1997) 722-726
6. Lio, S.C., Chen, J.S., Chen, J.L.: A Web Psychiatric Diagnostic Supportive System. Asia Management journal, 5(4) (2001) 521-539
7. Jen, J.F.: The Impact and Development of Medical Supportive System. Journal of Informatics, 5 (1997) 1-11
8. Chang, T.C., Lee, J.D., Wu, S.J., Wang, N.W., Shih, C.H., Yang, M.J.: Development the of a Psychotic Disorders Diagnosis and Medication Consultation System, T.S.M.H Medical & Nursing Journal, 10(4) (2004) 185-191
9. Kruschwitz, U.: An Adaptable Search System for Collections of Partially Structured Documents, IEEE Intelligent Systems and Their Applications, 18(4) (2003) 44-52
10. Ahn, K.H., Kim, S.K., Hong, K.P., Han, K.J.: Design and Implementation of Browser/Server Environment-based Hospital Information Search System (BS-HISS), TENCON 99. Proceedings of the IEEE Region 10 Conference, 2, (1999) 1569 – 1572
11. Ministry of Interior: The Revised the Standard Chart of Assistive Devices for Clients with Disabilities, (2004)

EASTIN: A Trans-national Information Network on Assistive Technologies

Renzo Andrich[1], Roberto Da Dalt[2], Renzo Giust[2], and Andrea Agnoletto[2]

[1] Centro di Bioingegneria Fondazione Don Carlo Gnocchi Onlus, Milano Italy
renzo.andrich@siva.it
[2] Don Gnocchi Sistemi, Milano Italy

Abstract. EASTIN (European Assistive Technology Information Network) is a trans-national information service on Assistive Technologies for people with disabilities, that aggregates the contents of six national information systems in Italy (Portale SIVA), Germany (Rehadat), Denmark (HMI Basen), the United Kingdom (DLF Data), The Netherlands (Hulpmiddelenwijzer) and Spain (Catálogo de Ayudas Técnicas). It is the result of the EASTIN project, carried out in 2004-2005 with partial funding by the EU within the *eTEN* programme. In order to work as a network, the six national systems have been *harmonized* according to commonly agreed standards and *integrated* through the new *EASTIN website*. This is equipped with advanced search engines able to perform search and retrieval operations across all the partners' databases, in any of the partners' languages.

1 Background

Information plays a key role in Assistive Technology (AT) service delivery systems. It helps empower people with disabilities and their families to make informed and responsible choices of AT devices; it is required by health care professionals when assessing their clients' needs, recommending appropriate AT solutions, training clients in their use, and designing rehabilitation, education or social participation programmes; it is vital for AT suppliers and manufacturers to better know the market, to discover opportunities, to find out ideas for development, to disseminate awareness on their products; it is important for policy makers and officers involved in public service delivery systems (insurances, Health Authorities etc.) in order to efficiently allocate resources in AT provision; people working in research and development also need access to qualified information that helps know what already exists, what users' needs are still unmet, what areas in the AT domain are admitting of significant developments.

In many Countries national information systems have been created to respond to these information needs. Some of them have long history, such as the Italian SIVA (started 1981), the English DLF-Data and the German Rehadat.

The importance of ensuring adequate AT information to citizens was acknowledged by the EU several times [1]. A major EU initiative in this field was the *Handynet* project (1988-1996) where an attempt was made to gather all data from the national systems into a common European multilingual database. The *Handynet*

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 524 – 531, 2006.

database was released in various Cd-Rom editions from 1993 to 1997. Although *Handynet* was unable to take wing as a self-supported product – and the project was greatly criticised because of the huge financial investment involved – it generated a set of standards and a common thinking that were taken up later by most national systems, both those that were in existence at that time and those that were established later.

This heritage was certainly one of the factors that made it possible – years later – to think about the possibility to create a trans-European information system on AT. The other factor was of course the move of all systems towards the Internet. This new technical environment offers new avenues for international integration of information, in terms of a networking systems rather than building new databases.

In 2003 the institutes responsible for 7 major information systems throughout the world founded – through the signature of a *Memorandum of Understanding* - the *International Alliance of Assistive Technology Information Providers*. In such “Memorandum” the members committed themselves to work together as a network in order to capitalize on each others’ experience, improve the content of all information systems, extend the user base internationally, achieve best practice, offer the best service to the systems’ users. The Alliance (www.ati-alliance.net) now includes 9 partners.

EASTIN – which stands for *European ASsistive Technology Information Network* – was the first concrete project generated within the Alliance. It took shape within the framework of eTEN, a EU programme designed to help the deployment of telecommunication networks based services (e-services) with a trans-European dimension.

2 Overall Design of the EASTIN Network

The six information systems taking part in the EASTIN networks are:

- *SIVA Portal* (`www.portale.siva.it`) run by Fondazione Don Carlo Gnocchi Onlus on behalf of the Italian Ministry of Welfare (in Italian and English)
- *Rehadat* (`www.rehadat.de`) run by the Institut der Deutschen Wirtshaft Köln, Germany (in German and English)
- *HMI-Basen* (`www.hmi-basen.dk`) run by the Danish Centre for Technical Aids in Tåstrup, Denmark (in Danish)
- *DLF-Data* (`www.dlf.co.uk`) run by the Disabled Living Foundation in London, UK (in English)
- *Hulpmiddelenwijzer* (`www.hulpmiddelenwijzer.nl`) run by iRv in Hoensbroek, The Netherlands (in Dutch)
- *Catalogo de Ayudas Técnicas* (`www.catalogo-ceapat.org`) run by Ceapat on behalf of the Spanish Ministry of Social Affairs (in Spanish).

Each EASTIN partner has a national role in his Country as producer of nation-wide databases and in some cases also of educational material that is widely distributed directly to citizens or through information networks. The funding of each system is based on national resources such as governmental agencies, pay-per-service schemes (mainly yearly subscriptions), sponsorships, or a mixture of the above schemes.

Before the EASTIN project started, such systems differed considerably from each other in terms of structure, technical approach and language. In the course of the project, they have been gradually *harmonized* (in terms of *consistency of scope, recognizability of items* and *coherence of contents)* and *integrated* (in terms of offering *a unique entry point* with *possibility of cross-databases searches*, and being managed through *a co-ordinated data collection process*).

A basic harmonisation requirement was about the type of contents to be handled by the network. This was categorised according to three major headings: *products information*, *companies information*, and *associated information*. The latter includes all information not strictly related to a specific AT product, but rather to "horizontal aspects" of this domain of knowledge such as *fact sheets*, *case studies*, *frequently asked questions* (FAQs), *ideas* and *links*: in other terms, to elements of knowledge that help grasp a higher lever understanding of the domain of assistive technology.

The integration / harmonisation process went through three validation stages with external user groups, respectively composed of *end-users*, *health care professionals* and *AT industries*.

Now the common website of the EASTIN network serves as system integrator providing unified *multilingual* access to all the systems taking part in the network, offering user friendly guidance and orientation, and providing Companies with a quick method to communicate with the administrators of the various national systems.

Fig. 1. The EASTIN homepage: `www.eastin.info`

3 How the EASTIN Network Works

After selecting the required language in the EASTIN homepage, a number of options are available such us getting information on the EASTIN activities, downloading documents, submitting information about new products, and cross search facilities across all databases. Here the search facilities are briefly described through a simple example.

Let's imagine we want to search for special computer keyboards and learn about their choice criteria. The *AT product search* function will find out for us all products of this kind available on the market of the six partner Countries.

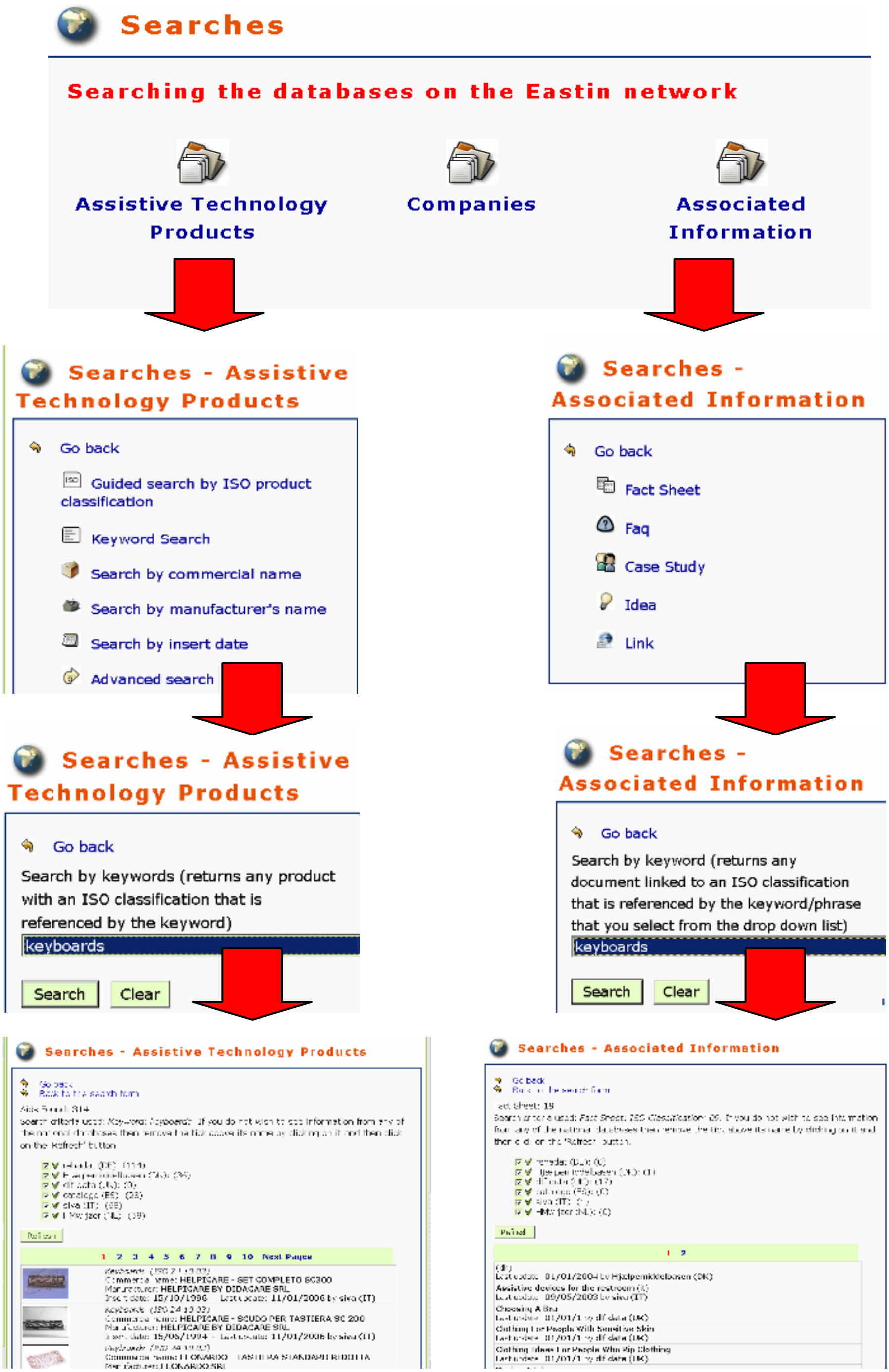


Fig. 2. Example of cross-databases search

Then we may wish to find out whether there are fact sheets in the *associated Information* providing some insight into choice criteria in relation to the user needs. Fig. 2 shows how these searches may work. In both cases the results lists show how many documents were found in each national databases. For each document, an extended record with full information can be shown in the required language.

4 Technical Approach

The technological approach adopted for the EASTIN search engines is the one commonly known as *web services*. This is a technique purposely developed to exchange data among different web-based information system.

The architecture involves two main set of components, respectively labelled *web service servers* and *web service clients*.

A *web service server* is a set of *functions* that can be called by referring to a web URL. Each *function* can receive one or more parameters and return a single result. The parameters and the result can be simple data types (e.g. string, integer, date,…) or complex data types (array, collection, object, …).

Web service clients are applications that can execute remote calls to functions belonging to one or more *web service servers*, by means of a protocol called SOAP (Simple Object Access Protocol).

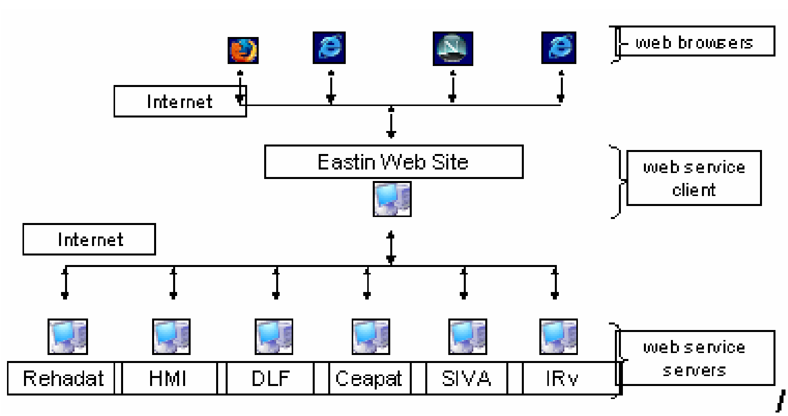


Fig. 3. The EASTIN webservices approach

Web services have a number of advantages over other possible methods that have similar purposes: they are fully platform-independent, they are language independent, they are suitable to in distributed databases applications, they introduce an abstraction level between clients sending requests and the RDBMS answering the queries.

The interface description is carried out via WSDL (Web Service Description Language). This is a XML standard for the description of the functionality, parameters and return values of the operations and of the interchange formats of a web service's messages. The EASTIN web service client is able to read the WSDL in order to determine which functions are available on the server. All particular data types in use are integrated in XML format in the WSDL data file. Now the client can use SOAP in order to call up a function that is listed in WSDL.

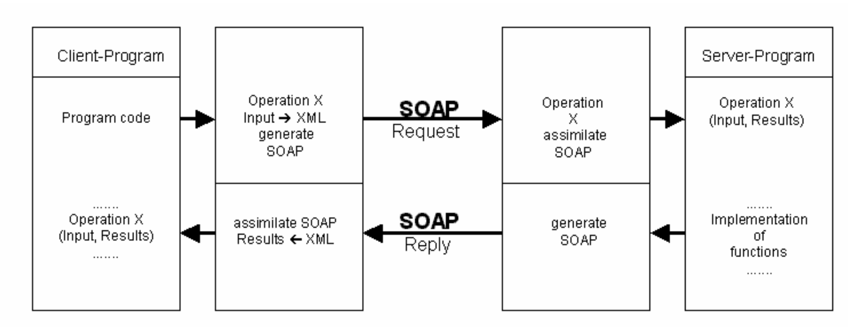


Fig. 4. Dialogue between the Eastin website (left) and the national information system (right)

5 The Multilingual Issue

All coded information (e.g. classifications of a product) is automatically presented in the required language on the basis of internally pre-defined translation tables. Conversely, information stored as free text is handled by software translation engines.

As the current state of the art of machine-translation is far from being perfect, the generation of understandable and usable translations of free-text documents was a great challenge. Comparative testing carried out within the EASTIN project over the available technologies – especially those who are able to perform not only *word-by-word* but also *sentence-by-sentence* translation – showed the possibility to achieve reliable understanding, provided:

- the translator is *trained* to deal with the specific language of the AT domain.
- native texts are written according to certain style rules.

Following to a number of tests, each partner chose an appropriate professional product able to translate from native language to English and vice versa, and trained it until achieving successful results. By running in batch mode, the translator generates English translations of each text as needed, so that users are able to retrieve any information in their language and in English from any EASTIN database.

As a further step, in the future – after the search results are generated – the user may be able to ask for a retro-translation into his/her language. Initially this will be done in batch mode (e.g. by getting the retro-translation later by e-mail); later, the possibility to have the text adequately translated on the spot in real time will be explored.

6 The Market Validation Exercise

In the course of the Eastin project, major efforts were devoted to the *market validation* of the new network, so as to ensure that the assumptions and the technical specification gradually developed would meet the real audience's expectations.

The validation activities were carried out through *external* assessment by a sample population representing the three major target groups of the EASTIN network (*end-users, professionals* and *industrialists*), and through *internal* self-assessment by the partners themselves.

For each target, a *core group* and an *extended group* of *external* validators were recruited in each Country. Each *core group* was composed of a small number of selected representatives who received specific training, and interacted directly with the Partners' project teams in all stages of the validation. Conversely, each *extended group* was composed of a larger number of people who candidated as validators, received no training, and interacted with the project team only online at the time of the validation exercises. These groups included both expert and novice Internet users so as to represent a meaningful sample of the "usual" population that will be confronted with the system. Overall, the number of validators accounted to 654.

Table 1. Population participating in the various validation rounds

External validators	Core groups	Extended groups
End users	n. 60 (10 / Country)	n. 300 (50 / Country)
Professionals	n. 30 (5 / Country)	n. 180 (30 / Country)
Industrialists	n. 24 (4 / Country)	n. 60 (10 / Country)

As the network developed gradually in the course of the project, the market validation activities were carried out in three rounds.

The *1st validation round* (May 2004) addressed the following question:

- *What are the weaknesses and the strengths of the six systems?*

This round was based on the *delphi method* and analysed five indicators, labelled as "the five As": *accessibility*, *availability*, *awareness*, *appropriateness* and *affordability* [2].

The *2nd validation round* (December 2004) addressed the following questions:

- *To what extent do the six systems deserve to be harmonised and integrated into a trans-European network?*
- *What is the added value of the network Vs the national systems?*

To this end, the same core groups were requested to validate a set of harmonisation and integration requirements as proposed by the Eastin partners. The validation was based upon a *mock-up*, which simulated through an animated slides-show how the future EASTIN website would work as network integrator. Again, the methodology was based on the Delhi method administered through a on-line questionnaire.

The *3rd validation round* (May - September 2005) focused on the *EASTIN network as a whole*. It addressed the following questions:

- *Technical sustainability: does EASTIN work and perform adequately?*
- *Usage sustainability: how will EASTIN be used by its potential market? Will EASTIN meet the potential markets expectations?*
- *How will EASTIN be financially and organisationally sustainable?*

For the first two questions, both the *core* and the *extended* groups were involved. During the four months validation period, the "navigation behavior" within the Eastin network was monitored by means of *log-files*; at certain times, on-line questionnaires were administered to collect the validators' opinions; at the end, a number of face-to-face or telephonic interviews were also carried out to complete data collection.

In brief, the findings of this validation round provided the ground for:

- re-designing the Eastin website and its search facilities, in order to better meet the audience's expectations and motivations
- taking decisions on the deployment strategy, including the organisational / financial model for the sustainability of the network
- finding out what contents areas deserve priority investments in data collection, in order to better meet the audience's interest
- addressing awareness and advertisement efforts in the deployment stage.

7 The Way Forwards

In order to deploy, maintain and further develop the netwok, in the long run, an *International Association* has been established as a legal entity registered in Italy. The founding members are the Eastin partners.

Acknowledgements

The following people should deserve mention for their contribution in building up the EASTIN network: A.Caracciolo, L.Garavello, M.Bulgheroni (Fondazione Don Carlo Gnocchi Onlus, Milano IT); N.Penn-Symons, R.Turner, R.Harvey, B.Self, D.Clarke (Disabled Living Foundation, London UK); M.Selbach, P.Winkelmann, F.Tomaszewski (Rehadat, Köln DE); M.Holm, E.M.Hansen, T.Lyhne (Danish Centre for Technical Aids, Taastrup DK); H.Knops, M.Verdonschot, R.De Vrije (iRV, Hoensbroek NL); R.Noya, C.Rodriguez, C.Larraz, R.Zamorano (Ceapat, Madrid ES).

References

1. Access to Assistive Technology in the European Union, Employment & social affairs, Social security & social integration, European Commission, 2003
2. The Promise Checklist, Proceedings of the Promise Colloquium (Brussels, 27-28/10/1997). `www.stakes.fi/promise/colloq/prcocont.htm`

Participation in Development of Computers Helping People

Lars Bruhn, Jürgen Homann, and Gerlinde Renzelberg

Zentrum für Disability Studies, University of Hamburg
Sedanstraße 19, D-20146 Hamburg, Germany
zedis@uni-hamburg.de

Abstract. The frequently prescribed „Design for All" or „Universal Design" reduces the comprehension of participation of people with disabilities and their interests in technical development tendentious to a pure user involvement. In such case, the ethical implications of new technologies are not considered. In this article we develop a concept for participation that will be more than a pure user involvement. To this purpose, ways to participation from the viewpoint of Disability Studies as well as Technology Assessment will be combined. The Participatory Action Research, which is suggested as methodical approach to technological development process will also be embedded in this concept for participation.

1 Introduction

Human being is a defect creature, who needs tools not only to adapt to his environment but also for his environment to adapt to him. This anthropological knowledge might sound uninteresting, but it's not all that uninteresting to people with disabilities. It's important for people with disabilities that tools, which exclude their usage, be adapted to their possibilities. The concept of "Design for All" or "Universal Design" that adjusts to a maximum without barriers for the greatest number of users as possible is also responsible to insure that it's reasonable and economical to an aging society. Important words like "eAccessibility", "eInclusion" or "eQuality" are eloquent testimony to effective international development. Yet, the development of new technologies is always in the context of a particular society. There is a reciprocal effect between development direction of technical changes and the parameters of a modernising society. The technologies for inherent progressive purpose require possibilities to be permanently extended and improved, but are often not considered by designers and producers and thus is the neutrality of technological aims and instruments always in doubt. Conceptions of a "Design for All" / "Universal Design" must consider the aspect of consequential determinism (Ropohl 1979, 239) in form of technology inherent in the social, technical and philosophical discussions, in order to avoid disregarding and excluding the minority, who in their values and norms differentiate themselves from the hegemony.

2 Participation

There are many initiatives in the areas of development and improvement of Information and Communication Technologies as well as Assistive Technologies that through

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 532 – 535, 2006.

the support of active user involvement (e.g. Verelst 2004) or with the help of training Accessibility Champions (Tusler 2004) aimed at reaching a development without barriers will lead to a far-reaching participation. It is noticeable that with regard to the participation of people with disabilities in technological development processes, the focus is appropriate clearly to the group of users and hardly goes out of the range of user involvement. Accordingly ethical implications or contradicting interests of potential users, who reject the use of certain technologies because of fear and/or expected consequences, threaten to remain to a certain extent technologically unconsidered and consequently, be at worst socially excluded.

A good example is the Cochlear Implant (CI). The development of CI (since in the 1980's) has always been considered by the Deaf as a threat to their culture and language without any effect on the medical field (in particular the consultation during the decision-taking process) and the pedagogy of the deaf. Contrary, an opposite effect is produced to the emancipatory achievements of the Deaf and as a result of this, deaf is increasingly considered a deficit again (Campbell 2005).

2.1 Participation and Disability Studies

Disability Studies (DS) is a new political and interdisciplinary scientific approach that is adopted to a paradigm for social construction of disability. Accordingly the focus does not lie on individuals affected by disability, but on cultural, economic and social contexts, from which disability results. The most decisive issue is the perspective of people with disabilities. Only this way can succeed to get people with disabilities to be expert in issues that concern them and thus promoting and supporting emancipatory processes, which will lead to self-determination. Apart from the central meaning of the interests of people with disabilities it is above all crucial in DS that disability is understood not individual-theoretical as medical-biological impairment, but as charateristic of cultural and social differentiation (Waldschmidt 2003, 12f). Seelmann (2005, 9) also confirms that people with disabilities were not well represented in scientific domain and asks for DS to look for possibilities to integrate people with disabilities in the field of science and development.

2.2 Participation and Technology Assessment

Development and changes in technique always affect the whole society in their social and economical structures. To believe that technical systems are not to save more than the intended purpose, is described by Ropohl (1996, 214) as a symptom of incredible technological naivety. It must be rather seen that every development of modern technique is in a comprehensive technological, economical, social, political and cultural context. To that extent the development direction and dynamic of new technologies will remain embedded into the structures of the respective society. The consequences and effects of technological improvement strongly influence the human way of life and also the cultural image of a society, which is why it is very important to continuously assess technique with help of empirical Technology Assessment (TA) in development stages in order to be able to react appropriately to unforeseen consequences. Revermann (1998, 16) therefore decided that the purpose of TA is to put in place an empirical foundation for the judgement, description and analysis of development, application and the potential consequences of a technique. Technical analysis in this sense includes judgement and analysis by the society. Correspondingly, ethics and TA are supposed to be in a complementary relationship to one another, where both have the task to make material information as well as normative orientation available. The only difference between ethics and TA is the way their programs are accentuated. While the priority of ethics is based on assessment of technique related means and purposes with the help of moral principles, the priority of TA lies on empirical analysis, descriptive and prospective assessment of potential consequences.

As far as TA is understood from ethical viewpoint as dialog between representatives of controversial positions, it is participative. As far as TA is orientated to problems induced by technologies, the technology does not stand itself unanalyzable at the beginning of a technological development, but the formulation of a specific demand is the condition for (further) technological development. That has to be negotiated in a discourse (with open results), without an obligation to consensus, but to make comparison orientated to problems of contentious development and alternative actions possible (Wevelsiep 1999).

2.3 Participatory Action Research

Action research is a suitable approach for different groups or people with the same aim of solving problems in development processes and thus improving development. According to Cunningham a problem is the definition of a need for changes and describes ways, how to deal with certain questions (Hart/Bond 2001, 62). The aim of action research is to broaden the competence of those taking part in a research process and to help those concerned to emancipate from the stipulated contents and from the experts' culture. Action research is applied above all in the building of an appropriate organizational infrastructure (a.a. O., 15).

The understanding developed from DS and TA is well integrated in Participatory Action Research (PAR). PAR is defined by the National Institute on Disability and Rehabilitation Research as a research strategy to involve those concerned in all stages of creation and research application (Doe/Whyte 1995, 2).

Action research is a research strategy, which cycle is very similar to that of development process of new technologies. PAR includes those affected by such development and their sometimes contradictory interests in the development process to help working against potential problems as early as possible.

3 Conclusion

Each technological development is carried out in a social context, from which it cannot develop separately and be judged from ethical point of view. The involvement of users with disabilities in the control of development processes is an important contribution to guarantee Designs for All. The fulfilment of participation to be seen here, might be technological sufficient, but having a look at the human society as social connection with conflicting interests, it cannot be considered enough. Real participation from our point of view is one that considers beside regulative Design for All, the contradictory interests (due to cultural, social or other reasons) associated to technology partly or fundamentally. Technological development should therefore – accompanied by DS – be embedded in a PAR process that supplements user involvement in the dimension of a participative TA.

References

1. Campbell, F.: Selling the Cochlear Implant. In: Disability Studies Quaterly, Vol. 25. (2005).
2. Doe, T., Whyte, J.: Participatory Action Research. Paper presented at the National Institute on Disability and Rehabilitation Research Conference „Forging Collaborative Partnerships in the Study of Disability". Washington D.C. (1995).
3. Hart, E., Bond, M.: Aktionsforschung. Handbuch für Pflege-, Gesundheits- und Sozialberufe. Bern Göttingen Toronto Seattle (2001).
4. Revermann, Chr.; Was heißt hier Bioethik? In: Büro für Technikfolgen-Abschätzung beim Deutschen Bundestag, Brief Nr. 15. Bonn (1998) 9-16.
5. Ropohl, G.: Ethik und Technikbewertung. Frankfurt a.M. (1996).
6. Seelman, K. D.; Universal Design an Orphan Technology: Do We Need Both? In: Disability Studies Quaterly, Vol. 25. (2005).
7. Tusler, A.: How to Make Technology Work. A Study of Best Practice in United States Electronic and Information Technology Companies. In: Disability Studies Quaterly, Vol. 25. (2005)
8. Verelst, T.: Design for All. Experience-Based Strategies for User Involvement in IST Research Projects. In: Klaus Miesenberger et al. (Ed.): ICCHP 2004. Berlin Heidelberg (2004) 104-109.
9. Waldschmidt, A.: "Behinderung" neu denken. Kulturwissenschaftliche Perspektiven der Disability Studies. In: Waldschmidt, A. (Ed.): Kulturwissenschaftliche Perspektiven der Disability Studies. Kassel, (2003) 11-22.
10. Wevelsiep, Chr.: Reflexive Bio-Politik. In: BEHINDERTENPÄDAGOGIK, 2/99. (1999) 214-223.

Developing Pedagogical Multimedia Resources Targeting Children with Special Educational Needs

Paloma Cantón[2], Ángel Lucas González[1], Gonzalo Mariscal[3], and Carlos Ruiz[3]

[1] Facultad de Informática, Universidad Politécnica de Madrid
Campus de Montegancedo, s/n. Boadilla del Monte - 28660, Madrid – Spain
agonzalez@fi.upm.es
[2] Consejería de Educación de la Comunidad de Madrid
pcanton@educa.madrid.org
[3] SETIAM – CETTICO, Fundación General UPM.
Campus de Montegancedo, s/n. Boadilla del Monte - 28660, Madrid - Spain
{gmariscal, cruiz}@cettico.fi.upm.es

Abstract. An educational resource targeting children with learning difficulties should be based on special teaching aids to achieve its educational purposes. Teaching supported by new technologies should guarantee the accessibility of the contents introduced over the web, as well as through the use of software applications. In this paper, we present a methodology for creating such resources and its application to the Proyecto Aprender (Learn Project). This project involves developing an accessible web site targeting teachers and the public (parents, families, researchers…) generally, and developing an accessible application targeting learners in particular. The whole process of design, preparation and construction is conceived to meet learners' special educational needs.

1 Introduction

In today's society, information and communications technology (ICT) use is a regular part of most people's lives. In the field of education, there more and more tools making use of new technologies are being released to support teaching. Regrettably, however, not all these systems take into account the minimum requirements for creating software or a web site that is accessible for people with disabilities. This is even more important when the tool targets children with special educational needs, whose education will influence their future social integration.

In this paper, we present a methodology for creating resources of this type and its application to the *Proyecto Aprender*. This project was developed for the Spanish Ministry of Education and Science. It involved developing an accessible online resource targeting children with special educational needs. It includes an alternative curriculum adapted to learners' needs.

The resource is composed of a web site and an application. The web site is divided in three sections: teachers, parents and learners. The learners section includes multimedia activities organized by learning objects: skills, difficulty levels and real-life situations. This is an alternative and very stimulating way of supporting special educational needs. To make these activities accessible a number of software tools have

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 536 – 543, 2006.

had to be developed to adapt the activities for learners with a range of disabilities (deafness, blindness, physical disorders, learning difficulties…).

The remainder of the introduction indicates the points that motivated the creation of the resource and what legal requirements an accessible web site and an accessible software should meet. The following section lists a number of works related to this area. Section 3 defines the proposed methodology and its application to the *Proyecto Aprender*. Finally, the conclusions and future lines of research are defined.

1.1 Motivation

There is a sizeable population of learners with special educational needs who, precisely because of their special needs, have difficulty in accessing a standard curriculum, and curricular adaptations have always been limited.

ICT use is an opportunity for designing and providing adapted resources that can have a major educational impact. Additionally, the development of a tool like this should be based on a methodology, including guidelines for educators, graphic designers and technical developers and on how to coordinate the different professional profiles. Also it is important to organize contents in a common framework.

Nowadays, there are few resources available to professionals working in the education of learners with special educational needs. Furthermore, there is no comprehensive tool targeting a broad-based group of people with special educational needs.

1.2 Accessibility-Related Legislation

For a web site whose primary users are children with special educational needs, accessibility plays a key role for preventing exclusion from access to information. This concept is defined in the eEurope 2002 Action Plan [1] as “info-exclusion”, and there is now talk about eInclusion and eAccesibility [2].

As regards worldwide legislation and international standards related to information society accessibility, several standards have been published since the '90s and the legislators of many countries, including the European Union, have passed specific laws on information society accessibility.

AENOR [3] standardised the first standard in the whole world concerning the creation of accessible web pages. This standard was revised and extended, and divided into standards UNE 139802:2003 (“Computer accessibility requirements. Software”) and UNE 139803:2004 (“Web Contents Accessibility Requirements”).

Additionally, although they are not an actual standard, the European Union has decided to consider the Accessibility Guidelines produced by the W3C consortium through the Web Accessibility Initiative (WAI) [4] as *de facto* standards. These guidelines take into account accessibility for all.

The ISO/TS 16071:2002 [5] standard was developed more or less at the same time. It provides guidance on the design of accessible software.

Obviously, it is impossible to develop a resource that is accessible for 100% of users. For a project of this type, however, the resource should clearly be as accessible as possible. For the *Proyecto Aprender*, the following decisions were taken:

– Website requirements: Spanish law [6] stipulates that government web sites must be accessible. They are obliged therefore to comply with the W3C WAI WCAG

'Double-A' conformance level. Because of this, the Main and Secondary web pages (Home, Teachers, Public, Worlds and Scenarios) of the *Proyecto Aprender* website should be 'WAI-AA' compliant, as well as conforming to UNE standard 139803:2004 (Spanish standard that refers to web accessibility). With respect to 'WAI-AAA' rules, these will be met whenever possible.

– Multimedia application requirements: The multimedia activities in the Learners section are included in an embedded application (software). Therefore, the multimedia activities should conform to UNE 139802:2003 (Spanish standard) and ISO TS 16071-2003 (international standard), which refer to software accessibility.

2 Related Work

The use of new technologies to create educational contents for children with special educational needs is a boom area because they are proving to be a valid resource that appeals to and motivates these learners. Consequently, initiatives have been launched to create multimedia tools as aids for educating people with special needs, as well as easy-to-access and user-friendly Internet gateways and resources.

The Research and Treatment in Mental Health and Social Services (INTRAS) Foundation is a non-profit organisation that developed the AIRE computer programs [7] and their successor, GRADIOR [8], a multimedia neurophysiologic assessment and rehabilitation system that is highly flexible and adaptable to different types of users depending on their disability.

The "Hércules y Jiló" educational software was developed at the University of Brasilia (Brazil) [9]. This software targets learners with moderate intellectual impairment. It aims to stimulate the child by getting it to interact between real and virtual worlds.

CompuThera is a multimedia software application targeting children that offers a step-wise graded learning-to-read method. The system is designed to behave as a tutor and help children who have problems with traditional learning and need constant attention from educators [10].

SoftTouch's multimedia software "My Own Bookshelf" [11] is a system for creating, editing and customising interactive multimedia books starring by child's picture and creating books for a range of purposes, from learning to read to basic social activities. The "Running Start Books - Social Scripts" [12] have been developed precisely in this area. These books aim to teach children with autism nine different social activities, like "When I am Mad", "Don't Hit Myself", "How Do I Say Hello?", etc.

SEN-IST-NET, the Information Society Technologies for Special Educational Needs (IST-SEN) gateway [13], was set up with the aim of compiling all the available resources and providing a platform for discussion and debate. This is a resource and publications guide on educating people with special needs, but it is also a European forum aiming to set coming trends in this field.

In the field of education, there more and more tools making use of new technologies are being released to support teaching [14]. However, none of the tools listed above complies with all the specified requirements. The *Proyecto Aprender* complies the requirements: it supports the learning and teaching of children with a wide range of special needs, it supports social integration and it provides guidance for teachers and parents on the education of children with special needs.

3 Methodology for Developing Contents

This section presents the different disability types, justifies the choice of contents (learning objects) that the resource will include, describes what tools and adaptability systems should be associated with each disability and, finally, explains the process of generating multimedia activities based on these particulars. To illustrate how these ideas were applied, we present the Proyecto Aprender, produced for the Spanish Ministry of Education and Science.

3.1 Proyecto Aprender

This was a two-year project, forming part of a more ambitious programme, the Internet at School Programme, composed of 14 educational projects [15].

The Spanish Educational Quality Organic Act (LOCE) [16] classes learners with special educational needs as follows: mental, physical and sensory impairments and/or personality disorders. Mental impairment refers to learners who fit the definition of mental retardation proposed by the American Association on Mental Retardation (AAMR) in [17]. According to this definition, an individual's strengths and limitations are present in each of these five dimensions: intellectual abilities, adaptive behaviour (conceptual, social and practical skills), social participation, interactions and roles, health (physical and mental health) and context (environments and culture).

This definition also identifies what sort of supports the person needs. These supports are resources and strategies that aim to promote the development, education, interests and personal wellbeing of the person and improve his or her individual functioning. Alternative behavioural programmes and the adaptive skills curriculum are the specification and development of the support activities to be developed [18]. They mostly cover the adaptive skills areas proposed by the AAMR [17].

The preparation of these support materials to plan and structure non-academic contents is costly in terms of time and input, which is the chief explanation for the huge shortage of support materials for professionals working in the fields of education, mental health and social services. The *Proyecto Aprender* directly connects with the development of support materials, like alternative behavioural programmes and the adaptive skills curriculum, closely related to life skills programmes [18].

In the development of this project we have taken into account that the target learners have special needs insofar as their ages range from 5-6 years all the way up to 17-18 years. Similarly, knowledge or curriculum competency levels will also vary enormously, and we are likely to encounter learners with slight-moderate mental retardation and others with severe mental retardation.

The overall objective is to build up and develop the physical, emotional, cognitive and communicational skills of learners with special educational needs, using new information and communications technologies to promote, to the greatest possible extent, their personal autonomy and social integration. Other specific objectives are: (a) establish multimedia activities related to personal autonomy, everyday problem solving and decision making, (b) learn to live in society through the knowledge of rules of cooperation and participation, (c) develop understanding, knowledge, linguistic skills, memory, logical reasoning and everyday problem-solving.

The learning objects are presented as metaphors (worlds, scenarios, characters, objects) to make the contents more appealing. Each world is a learning unit (see Fig. 1).

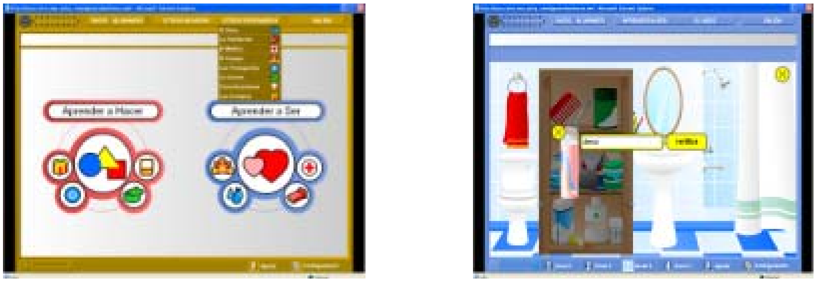


Fig. 1. Left: Menu and Worlds Screen, Right: Learn Activity Screen

Each module is composed of several learning objects that develop this module:

- Module 1. Learning to be. *The human body* (Evolutionary changes, Self-esteem). *Personal hygiene* (Routines, see Fig. 2). *Clothing* (Tidiness, Proper use). *Health* (Illness, Healthy habits)
- Module 2. Learning to live together. (To be developed) *At home. Games and sports* (Leisure). *Festivities and entertainment* (Leisure). *Rules of courtesy.*
- Module 3. Learning to do. *Scenario: Shopping and money* (Handling money). *Transport*. *Food* (Healthy diet, Preparing simple meals). *The media* (Press, Television, Computers).
- Module 4. Learning to know (this *world* is under construction).

The learning objects should meet three requirements: (a) take no longer than a maximum of 20 minutes or several similar fractions of time to do (for optimum application in a 45-minute classroom session), (b) contain an instructional guide including the activities to be performed in each learning object and explain the purpose of the resource and the objectives pursued by each interaction, (c) include an assessment to be set by teachers or a user self-assessment.

Four possible learning levels have also been established: (1) stimulation level; (2) infant (pre-primary) education level, (3) 1st-, 2nd-, 3rd and 4th-year primary level, and (4) 1st- and 2nd-year lower secondary level.

3.2 Adaptability Tools

It was considered necessary to implement new usability and additional adaptability systems to what were already provided by the tool used to implement the activities (Macromedia Flash MX 2004): Navigation Framework (usefull for *All* disabilities), Subtitles Bar (*Deaf* and *LD* generally), Manual and Automatic Scanning (*Physical Impairment*), Speech System (*Blind* and *LD* generally), Alternative Descriptive Texts (*Blind*), Text and Picture Sizes (*Visual Impairment*), Accessibility Options Settings (*All*). Additionally, certain pedagogical principles have been followed as regards text format -use of capital letters at low levels, language adapted to each level…- and graphics -colour contrasts, styles, layouts...- (LD generally).

3.3 Methodology

The development of a project of this magnitude implies the integration of different professional profiles that have clearly defined tasks and responsibilities. Apart from the coordinator of the whole team, three different profiles were established for the *Proyecto Aprender*: ***contents developers*** are pedagogues and educators, whose mission is to prepare the contents; ***graphics developers*** produce the graphics and audio-visual material; and ***technical developers*** are software engineers and programmers who develop the new adaptive tools, design web pages and program the activities.

In turn each profile has a manager (Contents manager, Graphics developers manager and Technical developers manager, respectively). The communication between the different groups of professionals is fundamental for the success of a project of this type. A contents developer must consult graphic designers and technical developers about the complexity of implementing selected functionalities within an activity and seek advice about alternative methods as appropriate. Likewise, the contents developers should validate the work of graphic designers and developers to check that the activities conform to the specified requirements.

In the following, we describe the tasks that need to be performed to successfully implement an accessible multimedia activity concerning a learning concept and the associated documentation, as well as the participants in each task:

1. Learning Object Justification and Learning Object Goals (manager: Contents).
2. Special Needs Object targeted by the activity (manager: Contents).
3. Proposed activities: multimedia script (manager: Contents).
4. Detailed activity description: Story Board (manager: Contents; other participants: Graphics and Technical Developers).
5. Learning Object Evaluation (manager: Contents).
6. Instructions for Parents/Families/Teachers. Similar activities to reinforce work on the learning object (manager: Contents)
7. Graphic design (manager: Graphic Designers; other participants: Contents, Technical Developers)
8. Activity Implementation and Monitoring (manager: Contents; other participants: Graphics and Technical Developers)

An activity monitoring table (see Table 1. Activity Monitoring) was designed to monitor the development of an activity (from the technical viewpoint). This table

Table 1. Activity Monitoring

	Activity	Actor
	Construct Multimedia Script	Contents Manager
Checkpoint 1	Multimedia Script Checkpoint	Coordinator
	Story Board Construction	Contents Manager
	Story Board Correction	Graphics & Technical Managers
	Story Board Acceptance	Contents Manager
Checkpoint 2	Story Board Checkpoint	Contents Manager
	Materials Files Construction	Graphics Manager
	Materials Construction	Graphics Manager
Checkpoint 3	Materials Checkpoint	Graphics Manager
	Construction of new requested services	Technical Manager
	Resource development and integration	Technical Manager
Checkpoint 4	Implementation Checkpoint	Technical Manager
	Operational testing	Coordinator
Checkpoint 5	Validation Checkpoint	Coordinator

covers whole the process starting from the multimedia script that describes an activity through to its implementation. This monitoring process with several checkpoints improves activity validation, materials reuse and knowledge at all times of who should step in to solve a problem, that is, it enhances the coordination among group members.

4 Conclusions and Future Lines of Research

The research presented and its application to the *Proyecto Aprender* have resulted in the first public online educational resource, with a special curriculum targeting learners with special needs for all levels of Spanish primary and secondary education.

It includes a comprehensive set of learning objects, which has meant that the resource has had to be carefully organised according to different criteria: ‘learning objects’, ‘life skills development’ and ‘complexity levels for each activity’. This has led to a navigation menu enabling several paths to access the multimedia activities.

Another innovative feature is the comprehensive information and materials supplied to provide parents, researchers, teachers and the public involved with learners with learning difficulties (the resource’s primary target users) with guidance and help.

One of the most difficult tasks was to assure that the resource was accessible for people with a range of and even combinations of impairments. On the one hand, as regards web content accessibility, we have followed the guidelines listed in [5] to conform to WAI-AA. Additionally, as regards the multimedia application, we have aimed to comply with the guidelines defined in standard ISO/TS 16071:2003. But, above all, both contents developers, and audiovisual producers and technical developers had to “think like the learners”. A group of pedagogues was appointed to test the resource. They have made use of real users and assistive technologies in the many tests run. The WAI Webxact [19] and HERA (Spanish assessor) guidelines evaluators have been used [20] to evaluate accessibility of web pages.

Certain children were considered as the target of this project is because, viewed from an integrative viewpoint, they have special educational needs, and they are going to need particular teaching aids to achieve their educational goals. Therefore, what will shape a child with special educational needs is not its impairment but the conditions affecting its personal development, and this justifies the provision of certain out-of-the-ordinary teaching aids or educational services, of which this project is part.

Module 4 will be completed in the near future, the web site and software will be translated into different languages, and research will be focused on defining a more general methodology for developing accessible multimedia educational web sites.

References

1. European Comission for the Feira European Council “eEurope 2002: An Information Society For All” Brussels (2000)
2. European Commission: Communication COM(2005) 425 “eAccessibility” Brussels (2005)
3. AENOR. `http://www.aenor.es`. (2006)
4. W3C (World Wide Web Consortium): Web Content Accessibility Guidelines 1.0. Available at: `http://www.w3.org/TR/WCAG10` (2005)

5. ISO (the International Organization for Standardization), ISO/TS 16071, “Ergonomics of human-system interaction -- Guidance on accessibility for human-computer interfaces” 14th plenary meeting of ISO/TC 159. Paris (2002)
6. BOE. Ley de Servicios de la Sociedad de la Información y de Comercio Electrónico (LSSICE). Ley 34/2002. Spain (11th July 2002)
7. Rodriguez, B., Franco, M.: Helping the Elderly Keep Mentally Active. An intelligent multimedia system, which helps the elderly to overcome intellectual and cognitive impairments, allows therapists to offer a higher quality of care for their patients. Esprit, the EU information technologies programme. IST 99, Helsinki (November 1999)
8. Franco Martín, M.A., Orihuela Villameriel, T., Bueno Aguado, Y., Monforte Porto, J.: Aplicación de Nuevas Tecnologías a la Rehabilitación Cognitiva por Ordenador. Programa GRADIOR. Revista Informática y Salud. Spain (1998)
9. Lacerda Santos, G.: Hércules y Jiló. Un Software Educativo para la Estimulación Matemática, Lingüística y Social de Niños con Deficiencia Mentas. Revista de Informática Educativa. Spain, February (2000) pp. 32 - 47
10. Moore M., Calvert S.: Brief report: vocabulary acquisition for children with autism: teacher or computer instruction. Journal of Autism and Developmental Disorders, Springer-Verlag. ISSN 0162-3257 (2000) pp. 359-362
11. SoftTouch. `http://www.softtouch.com/default.aspx`. (2006)
12. SoftTouch, Inc.: Running Start Books - Social Scripts (2005). Available at: `http://www.softtouch.com/productdisplay.aspx?productid=RS%20SS`
13. SEN-IST-NET. `http://www.senist.net`. (2006)
14. Sik-Lányi, C., Váry, Á, Sik, A., Nemetz, A., Geiszt, Z.: Multimedia Programs for Children with Hearing Difficulties. ICCHP 04. Springer-Verlag, ISSN 0302-9743 (2004) pp. 14-21
15. Gertrudix, M., Gálvez, M. C.: Desarrollo de contenidos educativos multimedia del Ministerio de Educación y Ciencia (CNICE). SIIE’04, Cáceres, Spain (2004)
16. BOE, Ley Orgánica de la Calidad de la Educación. Ley 10/2002 Spain (24.12.2002)
17. Borthwick-Duffy, S.: Review of "Mental Retardation: Definition, Classification, and systems of supports". American Journal of Mental Retardation 98 (1994) 541-544
18. Verdugo, M.A.: Programa de Habilidades Sociales (PHS). Programas Conductuales Alternativos. Amarú, Salamanca (1989/1997)
19. WebXACT (Watchfire). Available at: `http://webxact.watchfire.com`. (2005)
20. HERA. Available at: `http://www.sidar.org/recur/index.php` (2005)

Accessibility Issues in Two Specifications for E-Learning Tests: IMS QTI 1.2 and IMS QTI 2.0

Christophe Strobbe

Katholieke Universiteit Leuven – Dept. of Electrical Engineering,
Kasteelpark Arenberg 10, B-3001 Heverlee – Leuven, Belgium
christophe.strobbe@esat.kuleuven.be

Abstract. The IMS Global Learning Consortium developed the QTI (Question and Test Interoperability) specification to allow the exchange of question and test data, and their corresponding result reports, between learning systems. QTI 1.2 had some accessibility issues, as the VISUAL project discovered when transforming QTI tests into accessible HTML and voice user interfaces. Some problems were due to an insufficient mechanism to specify alternative text, other problems were due to the ambiguity of the intent of certain interaction types. QTI 2.0 solved the issue of alternative text, but the ambiguity with regard to the intent of interaction types was not sufficiently addressed.

1 Introduction

In Web accessibility, much attention is paid to accessibility guidelines for content, user agents and authoring tools, to software that checks content against accessibility guidelines, to the accessibility of user agents to assistive technologies, and to the development of assistive technologies and adapted devices. However, it is also important to review technical specifications of content types so as to check for accessibility features. If specifications are not reviewed for inherent accessibility issues, guidelines and tools can only work at the fringes of accessibility. The IMS Global Learning Consortium develops specifications for “distributed learning”, most of which are based on XML to facilitate the exchange of content, metadata etcetera between learning systems. One of these is IMS Question and Test Interoperability (QTI), a specification for tests and assessments, from single test questions (“items”) to sets of questions (“sections” and “assessments”), and interactive content. In addition to defining test questions and interaction, QTI files also contain “response processing” and grading information. QTI is an XML-based format that can be used to exchange tests and assessments between systems that use a proprietary format internally, but it can also used as the native format for tests in an e-learning system. At the time of writing, the current version is 2.0, which presents a thorough overhaul when compared to its predecessor, QTI 1.2. Version 2.1 is currently undergoing review.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 544 – 551, 2006.

2 QTI 1.2

2.1 QTI 1.2 Response Types

The QTI 1.2 specification consist of three documents, one if which is the *ASI Best Practice and Implementation Guide* [11]. This document lists 20 example basic item types. For each type, the guide provides a visual rendering and the corresponding XML code. For most of these types, it is possible to create an equivalent for use in a traditional class room, even though some of them cannot be realized with only pen and paper, for example standard multiple choice with audio (§ 4.1.4). The two types with sliders (multiple choice with slider rendering, § 4.1.8, and numerical entry with slider, § 4.4.3) are more typical of human-computer interfaces. Some item types, namely standard short answer (§ 4.3.3) and open ended questions with fill-in-blank (§ 4.3.1, e.g. "Name a renaissance playwright"), cannot be scored automatically with current technology.

The *ASI Best Practice & Implementation Guide* only provides a visual rendering of each of the items; the specification does not consider other renderings, in spite of the emergence of voice user interfaces, research on multimodal interfaces and the growing awareness of accessibility and inclusive design. From a pedagogic point of view, a number of examples would benefit from more extended feedback. The CETIS Assessment SIG worked on the examples from the specification and modified the XML where this was necessary or desirable[1].

The QTI 1.2 specification is agnostic with regard to the technology that is used to present the questions. Some online demonstrators use Flash for both the question text and the presentation of choices/input, others use a combination of HTML and Flash or Java applets. Flash and Java have accessibility issues, so the presentation implemented in the VISUAL project relied on HTML where possible and only resorted to an embedded interface where necessary. Nine out of the twenty examples in the specification can be handled with "pure" HTML, whereas the other eleven types require the use of embedded interfaces (at least, the examples in the *ASI Best Practice & Implementation Guide* suggest embedded interfaces).

The following types can be handled with pure HTML (paragraph numbers identify the sections in the *ASI Best Practice & Implementation Guide*):

- standard true/false (text: § 4.1.1);
- standard multiple choice (with text: § 4.1.2; with images: § 4.1.3);
- standard multiple response (text: § 4.1.5);
- standard fill in blank (text: § 4.3.1; decimal: § 4.4.1; integer: § 4.4.2);
- standard multiple fill-in-blank (text: § 4.3.2);
- standard short answer (text: § 4.3.3).

The following types require the use of embedded interfaces or DHTML:

- standard multiple choice (audio: § 4.1.4; with image hot spot rendering: § 4.1.6; with slider rendering: § 4.1.8);
- multiple response with image hot spot rendering (§ 4.1.7);

[1] These examples are no longer available on the Web.

- standard order objects (text: § 4.1.9; image: § 4.1.10);
- connect-the-points (logical identifier: § 4.1.11; XY response type: § 4.2.2);
- standard image hot spot (§ 4.2.1);
- numerical entry with slider (§ 4.4.3);
- drag-and-drop (images: § 4.5.1).

The above distinction is relevant to accessibility, because embedded interfaces have their own set of accessibility issues. For the first nine types, accessibility can be achieved by making sure that the HTML output conforms to the Web Content Accessibility Guidelines. However, this is not always possible with the QTI 1.2, especially with the examples provided in the specification. The following sections discuss these basic accessibility issues. These issues are excerpted from a report that was sent to the IMS Global Learning Consortium in June 2004.

2.2 QTI 1.2 Accessibility Issues

Place-Holder Text for Edit Boxes. Several response types require edit boxes in HTML: fill-in-blank is rendered with `<input type="text" …>`, and standard short answer is rendered with `<textarea rows="20" cols="80"></textarea>`. Checkpoint 10.4 of the Web Content Accessibility Guidelines recommends: "Until user agents handle empty controls correctly, include default, place-holding characters in edit boxes and text areas." User agents or assistive technologies that don't recognize empty form controls have become very rare (for example, JAWS 3.5 with Netscape 4.7.x), so it is probably no longer necessary to provide place-holder text. However, this was not the case when work on QTI began, and the QTI specification provided no information about this. All example basic item types use empty `response_label` elements, instead of using child elements in `response_label` for place-holding text.

Alternative text. The basic item type "standard multiple choice (images)" (§ 4.1.3) uses images; when rendering this item type in HTML, it is necessary to provide alternative text for accessibility.

The `altmaterial` element is an optional child element of the `material` element and is primarily intended to provide alternative language content in a single item. When used, there is one `altmaterial` element for each different language in the item; the `xml:lang` attribute indicates the language. The *ASI Information Model Specification* [12] explains that it is an element for "alternative content to be displayed in case the linked material cannot be rendered". A note explains that "this alternative material should not be of the same type as the original otherwise it too will not be rendered. Different versions should be used to support other languages or Accessibility options" (p. 42).

The following code sample adds alternative text to the first response label in one of the example from the *Best Practice and Implementation Guide* (mchc_ir_002.xml).

```
<response_label ident="A">
  <material>
    <matimage imagtype="image/gif"
      uri="mchc_ir_002_image1.gif"/>
```

```
      <altmaterial>
        <mattext>
          Round sign with white background, a red border…
        </mattext>
      </altmaterial>
    </material>
  </response_label>
```

The code sample shows that the relationship between the image and its alternative text is implicit. When the `material` element contains just one `matimage` followed by `altmaterial`, it is probably safe to assume that the `altmaterial` element contains alternative text for the image. However, there can be any number of `matimage` and `altmaterial` elements and there is no way to specify which `altmaterial`(s) go(es) with which `matimage`. This is different from the explicit links between media and alternative text (or between form fields and labels) that are used in HTML. For example,

- the alternative text for an image is provided as an attribute,
- the alternative text for an embedded object is provided in child elements within the `object` element,
- labels are explicitly linked with form controls by means of the `id` and `for` attributes.

The *Best Practice and Implementation Guide* should have stated more explicitly the importance of alternative text for accessibility. The DTD for QTI 1.2 should have made `altmaterial` and `mattext` required elements and should have specified a mechanism to "link" `altmaterial` with an image (or other media). In order to render an HTML `img` element with an empty `alt` attribute, implementers could leave the `mattext` element empty.

The same comments also apply to the use of other media (audio, video, …). The basic item type "standard multiple choice (audio)" (§ 4.1.4) uses sound files; when rendering this item type in HTML, it is necessary to provide alternative text for accessibility.

2.3 Accessibility and Test Validity

Alternative text in QTI would be necessary to allow "equivalent access". The *IMS Guidelines for Developing Accessible Learning Applications* state that *equivalent access* "provides the disabled user with content identical to that used by the non-disabled user", whereas *alternative access* "provides the disabled user with a learning activity that differs from the activity used by the non-disabled user" [2]. In the context of QTI, this distinction translates into "equivalent assessment" versus "alternative assessment". The challenge in alternative assessments is ensuring that one is assessing the same learning outcomes as with the original method; some assessments fail to do this [10]. Equivalent assessment should be provided whenever possible, but there may be tests where accessibility features conflict with validity constraints. Validity here refers to "[t]he degree to which accumulated evidence and theory support specific test scores entailed by proposed uses of a test" [1]. In some tests, accessibility features are essential for overcoming threats to validity, while in others, accessibility features can actually pose threats to validity [2]. Suppose, for example, that an art history test is intended to assess the student's ability to distinguish between specific chamber music genres (string trio, quartet, and quintet) and asks: "Which of the following extracts is played by a string quartet?" Providing text alternatives (for the hard of hearing, as required by WCAG [3] and the XML Accessibility Guidelines [4]) would threaten the validity of the test. In this case, one would provide alternative assessment to learners with hearing impairments; IMS has developed *IMS AccessForAll Metadata* [6] and *IMS Learner Information Package Accessibility for LIP* [9] to enable the retrieval of alternative content.

Some cases in the *Best Practice and Implementation Guide*, however, are ambiguous with regard to the aim of the question, and this makes it impossible to decide whether equivalent or an alternative assessment should be chosen. The example of drag-and-drop interaction contains a drawing of the solar system with an empty text box below each planet, and asks the subject to place text markers with the names of the planets inside the relevant boxes. However, it is not clear what the outcome is intended to test: whether the learner knows the order of the planets, whether the learner can recognize the planets by their relative size and colour, or something else. The intent of the test would determine if and what kind of text description of the planets would be appropriate. However, the relationship of this ambiguity with accessibility is not discussed in the QTI specification.

2.4 QTI and Voice Interaction

In the European project VISUAL, QTI items were not only transformed into accessible HTML (with XSLT): there were also efforts to adapt several interaction types to a voice user interface. Interaction with a voice user interface is different from the interaction of sighted users with a Web interface.

1. HTML specifies a two-dimensional layout, whereas a voice user interface works purely in the time dimension.
2. HTML is displayed in whole-page units, whereas VoiceXML and similar languages describe dialogues, which are in turn presented in smaller units (steps, forms, prompts, …).
3. An HTML page can present the user with dozens of options; voice applications must limit the number of options at any step in the dialogue to ease the burden on the user's memory and to improve the performance of speech recognition.

Generating voice interaction from QTI items is more problematic than generating a visual rendering, for several reasons. Since the QTI 1.2 specification only considers visual renderings, any rendering that is not purely visual is of necessity a proprietary extension of the specification (i.e. it overrides the presentation that is defined in the QTI specification). Moreover, the rendering format is often related to the didactic purpose of the exercise, so it may be inappropriate or even impossible (e.g. image hot spot/image map) to 'override' the rendering format with a different user interface. This problem can be solved by providing alternative Items with rendering formats that are adapted to the needs of the user. Voice interaction adds complexity to the process of answering a question. If the answer to a question goes beyond a single choice or input, the user should at any time be able to review the answer he has built up before completing it. Users should also be able to correct a partial answer before moving on:

this adds another level of complexity. Voice interaction is not only more complex than accessible HTML: there is also a greater difference between the interaction for interactive content or for assessments. (The dialogues that were modelled in VISUAL only considered interactive content. Assessments require that the user be able to review an answer and, if necessary, to correct it. With certain types of questions, corrections are very hard to handle in a voice user interface; it may even be easier to input the answer again from scratch.) Also, certain types of feedback which users of visual interfaces get 'for free' (e.g. maximum number of characters for text input) are not relevant to or can hardly be implemented in a voice user interface.

The complexity of voice user interfaces calls for some clarifications in the QTI specification. For example, QTI 1.2 has metadata to define

- whether or not feedback is to be made available (`qmd_feedback_permitted`), and
- whether or not hints are to be made available (`qmd_hints_permitted`),

but no metadata to specify if a question may be repeated or not when the learner does not immediately understand the question. This type of repetition is a feature of human-to-human communication ("I beg your pardon?", "Could you repeat that, please?") and needs to be handled in voice user interfaces. This is probably too fundamental for voice interaction to provide metadata or other means to disable this.

3 QTI 2.0

IMS QTI 2.0 (see [8] for an overview) is a complete overhaul of the language and allows authors to define the same types of questions (and others) with much leaner code. It specifies many types of interactions that were previously defined as extensions of the language (for example drag-and-drop to order items) or that were previously not defined (for example file upload and drawing). It uses many elements from HTML/XHTML, which makes transformation to HTML more straightforward. Reusing features from an existing language is usually good practice, especially from a language with documented and widely supported accessibility features (see checkpoint 2.9 in [4]). QTI 2.0 also borrows HTML's `object` element to allow authors to define alternative media and alternative text that are explicitly associated with each other (a suggestion from the feedback mentioned above).

Reusing HTML elements solved an important accessibility issue, but the new specification does not address the question of rendering in other modalities than visual interaction. The examples in the QTI 2.0 Implementation Guide still only consider visual renderings and the naming of most question types suggests a rendering instead of a learning outcome. The introduction to the implementation guide states: "[The screen shots] are designed to illustrate how a system might implement the specification and are not designed to be prescriptive. *Other types of rendering are equally valid*." However, the document does not discuss non-visual renderings, and the names of many interaction types suggest visual renderings (for example, `hotspotInteraction` and `drawingInteraction`). If the intent of the test is related to a specific visual rendering, it is often impossible to define an equivalent test without making the outcome invalid. However, for tests where the rendering is only a function of the learning system, QTI 2.0 does not define a mechanism that allows authors to specify that the rendering is not important. Moreover, the specification never discusses alternative renderings for question types that can be rendered in non-visual modalities.

QTI takes the opposite approach of W3C's XForms [5], the successor of HTML forms that will be integrated into XHTML 2.0. XForms are designed for accessibility because—among other reasons—they do not make assumptions about how forms will be rendered (for example visually or through voice interaction). When trying to convert the examples from the QTI 2.0 Implementation Guide into XForms, one would find that some of the interactions illustrated in the guide are ambiguous. The example for `associateInteraction` asks the learner to identify three pairs of rivals in a list of six Shakepearean characters. The illustration is a screen shot of a drag-and-drop rendering, and it is not clear if a rendering with edit boxes or selection lists would also be appropriate. In other words, it is not clear if the rendering is essential or whether one is allowed to render the question in other modalities.

4 Conclusions

The IMS specifications for Question and Test Interoperability, versions 1.2 and 2.0, have left certain aspects about the allowed renderings and the intent of questions undefined, and this leads to ambiguity with regard to how certain question types may be adapted for people with disabilities. If more guidance were provided on the relation between interaction types and the intent of the tests, it would be possible to create more equivalent assessments instead of always relying on alternative assessments. This would benefit people with disabilities directly, but it would also eliminate some of the overhead caused by alternative assessments and their associated metadata.

Acknowledgements

Part of this work (accessibility issues of QTI 1.2) was undertaken in the framework of the project VISUAL — IST-2001-32495 (October 2001-June 2004) — funded by the IST Programme of the European Commission.

References

1. AERA, APA & NCME: Standards for Educational and Psychological Testing. Washington DC: American Educational Research Association, American Psychological Associations, National Council on Measurement in Education (1999).
2. Barstow, C., McKell, M., Rothberg, M., Schmidt, C. (eds.): IMS Guidelines for Developing Accessible Learning Applications. White Paper. 27 June 2002. Available at: http://www.imsglobal.org/accessibility/accv1p0/imsacc_guidev1p0.html.
3. Chisholm, W., Vanderheiden, G., Jacobs, I. (eds.): Web Content Accessibility Guidelines 1.0 – W3C Recommendation 5-May-1999. Available at: http://www.w3.org/TR/WAI-WEBCONTENT
4. Dardailler, D., Palmer S. B., McCathieNevile C. (eds.): XML Accessibility Guidelines – W3C Working Draft 3 October 2002. Available at: http://www.w3.org/TR/2002/WD-xag-20021003

5. Dubinko, M., Klotz, L. L. Jr., Merrick, R., Raman T. V. (eds.): XForms 1.0 – W3C Recommendation 14 October 2003. Available at: http://www.w3.org/TR/2003/REC-xforms-20031014/
6. Jackl, A. (ed.): IMS AccessForAll Meta-data Best Practice and Implementation Guide. Version 1.0. Final Specification. 12 July 2004. Available at: http://www.imsglobal.org/acclip/acclipv1p0/imsacclip_bestv1p0.html
7. Lay, S. (ed.): IMS Question and Test Interoperability Implementation Guide. Version 2.0. Final specification. 24 January 2005. Available at: http://www.imsglobal.org/question/qti_v2p0/imsqti_implv2p0.html
8. Lay, S. (ed.): IMS Question and Test Interoperability Overview. Version 2.0. Final specification. 24 January 2005. Available at: http://www.imsglobal.org/question/qti_v2p0/imsqti_oviewv2p0.html
9. Norton, M., Treviranus, J. (eds.): IMS Learner Information Package Accessibility for LIP Best Practice and Implementation Guide. Version 1.0. Final Specification. 18 June 2003. Available at: http://www.imsglobal.org/acclip/acclipv1p0/imsacclip_bestv1p0.html
10. Sharpe, K., Earle, S.: Assessment, Disability and the Problem of Compensation. Assessment and Evaluation in Higher Education, 25. 2 (2000) 191-199.
11. Smythe, C., Shepherd, E., Brewer, L., Lay, S. (eds.): IMS Question and Test Interoperability: ASI Best Practice & Implementation Guide. Version 1.2. Final Specification. 11 February 2002. Available at: http://www.imsglobal.org/question/v1p2/imsqti_asi_bestv1p2.html
12. Smythe, C., Shepherd, E., Brewer, L., Lay, S. (eds.): IMS Question and Test Interoperability: ASI Information Model Specification. Version 1.2. Final Specification. 11 February 2002. Available at: http://www.imsglobal.org/question/v1p2/imsqti_asi_infov1p2.html

Informatics Teachers and Their Competences in Inclusive Education*

Ľudmila Jašková

Department of Informatics Education, Faculty of Mathematics, Physics and Informatics, Comenius University, Bratislava, Slovakia
jaskova@fmph.uniba.sk

Abstract. The paper deals about the problem of necessity to modify preparation of future informatics teachers so that they will be able to use information and communication technologies for their educational work with handicapped students in mainstream schools. In our research we try to find missing teacher competences and appropriate educational contents, methods and forms.

1 Introduction

Information and communication technologies (ICT) are necessary in many professions and also in everyday life. Independent life, further education and professional success of handicapped people are possible only if there are fulfilled these conditions:

- They can use modern information and communication technologies in their study and everyday life.
- They can study integrated in mainstream schools.

Many handicapped people study integrated in mainstream schools, but teachers at these schools are not prepared good enough to teach handicapped students and to use ICT in education. There is also luck of study materials and methodological materials for teaching ICT for handicapped students.

In frame of our planed research we are trying to answer these questions.

1. Why should be informatics teacher prepared for teaching handicapped students?
2. Which teacher competences are missing in former educational study of informatics provided by the Slovak universities?
3. How to help informatics teachers to be competent?

How to measure if informatics teachers are more competent after using our suggested methodology?

* This paper rose from the research realized in frame of the ESF project MIŠŠ 21, financed by the European Union. Európsky sociálny fond.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 552 – 559, 2006.

2 Why Should Be Informatics Teacher Prepared for Teaching Handicapped Students?

2.1 Legislation

Rights of children with special educational needs were reflected in the following documents [1].

- *The English and Welsh National Curriculum* which included "The entitlement to access to a curriculum which includes the National Curriculum and which is broad, balanced and relevant" (Department of Education and Science, 1988).
- *The Salamanca Statement* [7] of the UNESCO world conference on special needs education (June 1994) states that: Mainstream schools with this inclusive orientation are the most effective means of combating discriminatory attitudes, creating welcoming communities, building an inclusive society and achieving education for all.
- Educational provision for learners with special educational needs should be provided in the *„Least Restricted Environment"*(IDEA US, 1990).
- *Articles* 23 a 29 of the United Nations Convention *on rights of the child* (1989).

Within this context inclusion is irreversible process characterized by evolution rather then revolution. Schools can improve and become more effective at the same time as they become more inclusive.

2.2 Statistics

It is not easy to discover how many disabled people are living all over the world because not every handicapped person is registered. Information about disabilities is difficult to compare between different countries because different countries have different definitions of disability and different degrees of political will to publicize such information. The following statistics therefore relate primarily to the UK and to International Labor Organization sources [3].

- The ILO estimates that there are 610 million disabled people worldwide.
- Disability is estimated to affect 10 percent to 20 percent of every country's population, a percentage that is expected to grow because of poor health care and nutrition early in life, growing elderly populations and violent civil conflicts.

As we can see, each tenth inhabitant of Earth is disabled and it is not low amount. But it has to be mentioned that four hundred million of disabled people live in the developing countries and large amount of disabled is 50-60 years old.

2.3 Information and Communication Technologies

Technology is now the versatile tool for handling information and communication. Information and communication technology is therefore central tool for inclusion.

For learners with special educational needs, technology can provide them with access to the curriculum if appropriately utilized within their individual education plans.

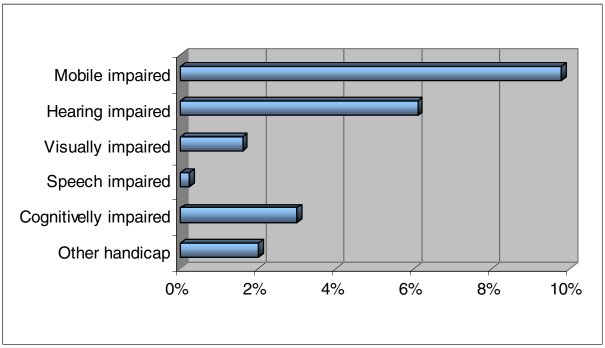


Fig. 1. Percentage of population in Europe with problems using information and communication technologies (in 2004) [4]

ICT can not only make education easier for handicapped students, but in integrated schools they are necessary. Forasmuch as informatics teachers are most familiar with ICT they should know how it can be used for handicapped students in education.

This necessity is forced also by the following roles of informatics teachers.

1. *Informatics teachers teach how to use ICT.* Whereas handicapped people use the same software as others**,** *they may use computers in different way.* When a teacher knows how to use computer and how to teach computer skills students without disabilities, this knowledge is unusable when he/she is teaching disabled students.
2. *Informatics teachers participate on developing school web pages and educational web pages. These web pages should be accessible for handicapped people.* When teaching development of web pages, informatics teachers should teach their students how to create accessible web pages for all.
3. *Informatics teachers are unofficially considered to be ICT-coordinators –general advisors for using ICT in education.* Therefore they should know:
   - how to use ICT for *transformation of study material* into accessible form for handicapped (e.g. visually impaired) students,
   - how can handicapped student *use ICT for taking notes, for his/her project work, for testing, and for communication with teachers and schoolmates*,
   - how can mainstream teachers and ordinary students *use ICT for discovering the way of living of their handicapped students and schoolmates* (where on the web it is possible to find information about handicaps, where are vision or hearing simulations, where to find information about Braille and sign language, etc).

3 Which Teacher Competences Are Missing in Former Educational Study of Informatics Provided by the Slovak Universities?

Our list of competences was created after studying the *following sources of information.*

- Educational standards and informatics curriculum for secondary schools published on the web page of the Slovak state pedagogical institute <www.statpedu.sk>.

- General competencies of teachers of visually impaired formulated during international workshops organized by European Committee of ICEVI. These workshops were for teachers/staff members of universities/colleges etc, whose duty it is to train the teachers of the visually impaired ([6]).
- Interviewing informatics teachers who have already taught handicapped students.
- Our own practice in teaching handicapped students.

Competencies – a specific range of skills, knowledge, or ability we divided into seven categories.

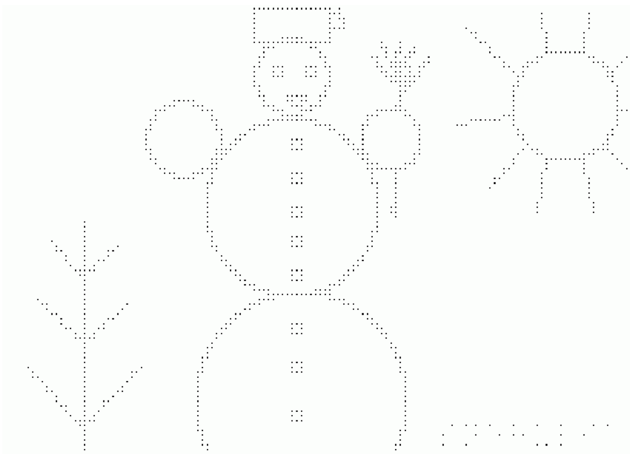


Fig. 2. Example of activity for sighted children - production of tactile picture for unsighted friend

Professional Competencies. The teacher should be able to demonstrate a know-ledge and understanding of:

- The appropriate information and communication technology and its applications for handicapped students.
- The principles of making accessible web pages and possibilities of how to use established networks to access resources.
- The possibilities of how to use and develop literacy and numeric skills and imaginations of handicapped students.

Psycho Didactical Competencies. The teacher should be able to demonstrate a knowledge and understanding of:

- The implications of impairment on physical, cognitive, emotional, social and language development and their effect in educational process.
- The mobility skills of handicapped students and possibilities of how to use them.
- Design and production of presentation and evaluation material in the appropriate medium for all students with handicap.
- The didactical principles necessary for teaching handicapped students.
- Ergonomic solutions for using computers.

Communicational Competencies. The teacher should be able to demonstrate a knowledge and understanding of:

- The appropriate methods of teaching and communication for children and young people who have handicap including those with multiple or dual sensory impairment.
- Use a range of interpersonal skills, appropriate for working with children, parents and professionals.

Organizational Competencies. The teacher should be able to demonstrate a knowledge and understanding of:

- Current national and local developments and legislation in curriculum and assessment and their implications for learners with special needs.
- Ability to design and manage an adapted teaching programme, taking account of the competing demands of the national curriculum and the special curriculum.

Diagnostic Competencies. The teacher should be able to demonstrate a knowledge and understanding of:

- The anatomy and physiology of the eye and ear and an understanding of visual and hearing functioning.
- The range and functions of support services available to those with a handicap.

Advisory Competencies. The teacher should be able to demonstrate a knowledge and understanding of:

- How to lead training sessions about ICT for handicapped for teachers, non-teaching assistants, parents and relevant others in the field of handicap.
- How to support the inclusion of a handicapped child or young person in the school and community using ICT.

Competencies of Self-reflection. The teacher should be able to demonstrate a knowledge and understanding of:

- How to monitor and evaluate the implementation of individual student programs and learning environments and how to plan the next stage in the education of the individual.

4 How to Help Informatics Teachers to Be Competent?

To improve competencies of informatics teachers it is necessary to make some changes in educational study of informatics for future teachers and to prepare courses for real mainstream teachers of informatics.

4.1 Modifications in Educational Study of Future Teachers

Regarding the list of competencies mentioned above and informatics curriculum for secondary schools we recommend to modify the study plan for educational study of informatics in the following way.

- To add the following topics to the existing subjects.
  - To add the topic about ways of using computers by handicapped students into the basic course aimed at gaining basic computer skills.
  - To add the topic about e-accessibility into the course on Development of web documents.
  - To add the topic about universal and accessible design into course on Development of educational software.
  - Production of educational software for handicapped children and computer games facilitating integration of handicapped students in frame of their term projects and diploma thesis.
  - Production of methodological material about teaching informatics for handicapped children in frame of student's diploma thesis.
- To provide a new course on ICT in education of handicapped students containing these topics [5].
  - Main advantages of using ICT in education of handicapped students.
  - Visually impaired people and special technical equipment for visually impaired people.
  - Computers for visually impaired – special SW and HW and methodology of teaching computer skills.
  - Hearing impaired people and special technical equipment for hearing impaired people.
  - Computers for hearing impaired – methodology of teaching computer skills.
  - People with mobility problems and their work with computer.
  - People with cognitive problems and their work with computer.
  - E-accessibility.
  - Adaptation of study material for handicapped students.
  - Principles of universal design for software development.
  - ICT in activities facilitating integration of handicapped students into mainstream schools.
  - Limitations of technology and problem situations.

4.2 Courses for Real Teachers

For the informatics teachers it is possible to gain the competences mentioned above attending these forms of courses.

1. Traditional course in computer room with special hardware for handicapped people.
2. On-line course with study materials published on the web. Communication between the teacher and participants would be via the Internet (using e-mail or video-conferencing or chatting software).
3. Development of web site containing the study and methodological material.

It is not possible to cover all the topics mentioned in part 4.1 in frame of one course. We suggest splitting the content into following modules.

1. Methodology of teaching MS Windows and MS Word for handicapped students.
2. Methodology of teaching Internet applications (e-mail, chat, WWW) for handicapped students.

3. Methodology of teaching MS Excel and MS Access for handicapped students.
4. Study materials and their adaptations for handicapped students.
5. Accessibility of information on web pages – legislation, development and testing.

Each module mentioned above would contain also this information.

- Who are handicapped students?
- Overview of special HW and SW for handicapped people.
- Didactical principles for teaching handicapped people.

5 How to Measure If Informatics Teachers Are More Competent After Using Our Suggested Methodology?

We plan to do a case study research in frame of which we would inquire these cases:

- experienced informatics teachers, who have already taught handicapped students (we will use structured interview),
- mainstream informatics teachers, who have never taught handicapped students before (we will use structured interview),
- handicapped students attending mainstream schools (we will use structured interview),
- students and teachers – attendants of our courses.

In frame of the research we plan to realize following forms of courses for future and real teachers.

- New course for students of educational study of informatics covering the topics mentioned in 4.1.
- On-line course for mainstream informatics teachers about e-accessibility.

We will write down the detailed information about the running courses – topics, used educational methods and forms, used study materials, feedback from students (immediate reactions, written homework). We will analyze all this material.

For collecting of data we plan to use also multiple choices questionnaires.

- Attendants will fill the same questionnaire before and after the course, we will evaluate separately data collected before and after the course and compare the results so that we will see how the competences of attendants have changed during the course.

Our research will be aimed at answering the following questions.

- Which competences of attendants were developed or improved during the course?
- Are the courses good enough for acquiring the necessary knowledge and skills?
- Are the study materials good enough?

We know that there are similar courses for informatics teachers in other countries, but the similar study have not been done yet. All the studies done before in the field of integrated education were aimed either at the handicapped students or at ICT for handicapped people. Our research is aimed primarily at informatics teachers and their knowledge and skills.

References

1. Blamires, M.: Enabling Technology for Inclusion. London: Paul Chapman Publishing Ltd, A SAGE Publications Company (1999) 194 p. ISBN 1-85396-394-1.
2. Campbell, L.F. et al.: Yes, they can!; Preparing Parents & Educators of Blind and Visually Impaired Individuals in the Field of Access Technology [online]. publication of EENAT project (2000) Accessible online: <http://www.lass.lt/en/kttc/modules/toc.htm>.
3. CSR Europe, Disability Facts and figures. Accessible online: <www.csreurope.org/csrinfo/csrdisability/DisabilityFactsandfigures/>.
4. GILL, J.: Access-Ability; Making technology more useable by people with disabilities [online]. London: RNIB. 2004. 40 p. ISBN 1 86048 030 6. Accessible online: <www.tiresias.org/guidelines>.
5. IKT vo vzdelávaní hendikepovaných – in Slovak, ICT in Education of Handicapped People, Accessible online: <www.edi.fmph.uniba.sk/jaskova/ikth>
6. International Council for Education of People with Visual Impairment, European Region. Accessible online: <http://www.icevi-europe.org/>
7. Salamanca Statement. Accessible online: <http://www.diseed.org.uk/Salamanca%20Statement.htm>
8. Section 508. Accessible online: <http://www.section508.gov/>
9. Web Content Accessibility Guidelines 1.0. Accessible online: <http://www.w3.org/TR/WCAG10/>

ECDL bf: Equal Opportunities Through Equal Access to an ECDL E-Learning Solution

Andrea Petz and Klaus Miesenberger

Institute Integriert Studieren, University of Linz, Altenbergerstraße 69, 4040 Linz, Austria
{andrea.petz, klaus.miesenberger}@jku.at
http://www.integriert-studieren.jku.at

Abstract. People with Specific Needs – even if they are, in principle, willing and capable of taking part in 'our' information - or better ICT - society – run the risk of unemployment and (social) exclusion significantly more often than people without disability. While recognized certification schemes, like the internationally recognised and promoted IT certificate ECDL help to clearly prove ones knowledge in handling mainstream IT and open up the labour market, those certificates build up new and insuperable barriers for people with specific needs when not designed and implemented accessible. Following the outcomes of the previously presented EU funded project ECDL PD, primarily accessible inclusive training settings are lacking. To overcome this obstacle, the Austrian project ECDL bf (ECDL without barriers) worked on an extensively accessible ECDL e-learning solution applicable for the mainstream market that will be presented in this paper.

1 Introduction

This paper presents our ongoing work in the Austrian national funded project 'ECDL bf'. This project is a national and in depth spin off of the EU (Leonardo da Vinci) funded pilot ECDL PD[1] (2001 – 2004) that aimed at adapting basic materials, identifying structural / organizational barriers for people with disabilities within the 'ECDL–Universe' and deploy lasting mechanisms to carry on the projects' efforts and keep the needs and requirements of people with disabilities on the agenda.

The follow–up ECDL bf started in late 2004 as a co-operation of strategic and operational partners out of economy, adult and continuing education (IT), research & development / higher education, governmental and non–governmental institutions.

Within the operational partnership, 'BIT Media and BIT Trainings Centre' is one of the best known ATEC in Austria, and the biggest supplier of training materials for ECDL oriented training; their materials were the basis for further developments. 'OCG' (Austrian Computer Society) drives the ECDL in Austria and is committed to keep the ECDL open for people with disabilities and is responsible for approving

[1] Please see http://ecdl-pd.aib.uni-linz.ac.at for further information.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 560 – 567, 2006.

teaching and training materials on behalf of the ECDL Foundation in Ireland following strict quality assurance guidelines.

Microsoft™ Austria took part and supported this project by funding (as did the Federal Bank of Austria and the Austrian Federal Ministry of education, science and culture) and disseminating our results having informal studies that stated most clients (users and training organisations) use Microsoft Windows and Microsoft Office products as the basis for ECDL [1].

1.1 User Needs Assessment and State of the Art

More than 35.000 people with disabilities in Austria were out of work in 2003 [2].

On the one hand, ICT is known to open up new job opportunities – in particular for well trained people on adapted workstations – independent of a possible disability. On the other hand, the outcomes of our already presented project ECDL PD stated clearly that there is a significant lack of accessible IT learning and training opportunities – especially those qualifying people with specific needs to enter (or return to) the open labour market. More precisely, there is a lack of accessible *and inclusive* IT learning and training opportunities, and specialized training settings often aggravate the situation. In most cases, qualifications earned in those settings are "specialized", too – and neither recognized nor accepted equally outside this special environment [2].

The project ECDL bf therefore headed to design and implement a comprehensively accessible ECDL teaching and training e-learning solution applicable for the mainstream market without changes to the originally designed accessible product.

The project focused on people with disabilities and elderly people suffering from age related disabilities. Statistics outline that 10 – 15 % of the population in western countries suffer from a disability and therefore have problems in using standard ICT.

Disability in the project's context has been defined as a possible constraint (or all possible combinations of constraints) of a persons' sense of sight, sense of hearing, sense of touch, mobility, cognitive and / or linguistic ability, spatial perception and finally perception and / or power of concentration.

Additionally, we identified the following barriers that take effect independent of a disability:

- Inaccessible information (concerning formats, usability…),
- Inaccessible communication (language, including technical terminology…),
- Inaccessible architectures / structures,
- Personal contact / support is not sufficient or not sufficiently adapted to a persons' needs.

These definitions emphasized the necessity of a comprehensively accessible setting and also showed a high potential impact on the mainstream market, so the work for ECDL bf started already in late 2004 [3].

2 Target Groups and Specific Needs

As the needs of different groups of people with disabilities seemed to vary very much and were evaluated as sometimes contradicting, it was necessary to define categories of user groups in order to be able to address them properly. In this project the following groups have been taken into account:

- Blind people / people with low vision and / or visual processing disorders,
- Deaf and hard of hearing persons,
- People with motor disabilities,
- People with specific learning difficulties.

All these groups do have problems in getting training as well as showing the skills in handling ICT [4]. The project therefore aimed to support better access to teching and training by developing training materials, ‘train the trainers’ initiatives and finally by initiating / supporting training activities.

2.1 Blind People

Blind people use their sense of touch and / or sense of hearing to access information on the screen. This means for the target group that

- Assistive technology is inalienable,
- Computer use is an active, tactile reading process running linear - letter by letter and is time-consuming,
- Blind users depend on additional information given by the screen reader.

Therefore, attention must be paid to:

- Clear and coherent structures,
- Comprehensive keyboard support (Shortcuts, Hotkeys),
- Equivalents for complex structures, such as pictures, graphs and tables.

2.2 People with Low Vision and / or Visual Processing Disorders

This target group accesses information on the screen using their senses of sight and hearing. This means in detail that:

- Assistive devices may be used,
- Flexibility and adaptability of the screen - output to the highest possible extent is necessary,
- Acquisition of information takes more time.

That implies that attention has to be paid to:

- Typographic framework (fonts, font size, font weight, kerning…) particularly suitable to ease the process of reading and comprehension without early fatigue,
- Readability / legibility support,
- Colour schemes and contrast (not only for people with achromatopsia but free customizable),
- Clear and coherent structures.

2.3 Deaf and Hard of Hearing People

Deaf people access information on the screen using their sense of sight. Hard of hearing people additionally use their sense of hearing to some extent.

That means for this target group that:

- Output and translations with subtitles and in Easy to Read are indispensable, gesture language (video) is highly appreciated,
- Visual output of auditory information (system / error messages, alarms…),
- Visual support (attraction of attention, support of the information displayed…),
- Readability / legibility support,
- Colour schemes and contrast in line with better legibility / readability and multimedia implementation,
- Typographic framework (fonts, font size, font weight, kerning…) particularly suitable to ease the process of reading and comprehension without early fatigue,
- Clear and coherent structures.

2.4 People with Specific Learning Difficulties

People with SPLD access information on the screen with the senses of sight and hearing. Because of the vast variety of possible forms of learning disabilities (and their combinations) the use of a 'common sense' is necessary to elaborate possible needs and requirements. Experience showed that this target group could benefit from materials adapted to deaf and hard of hearing persons enriched with auditory information and reader support.

2.5 Older Adults

This target group and their needs may be characterized by a possible combination of multiple, moderate constraints. Often there is no need for specialized assistive technologies, but the ability to see diminishes for example, as the ability to hear and keep focussed does. Perhaps there is also a constraint concerning the individuals' mobility.

2.6 People with Motor Disabilities

People with motor disabilities access information on the screen with senses of sight and hearing and there are no 'critical' guidelines to be followed concerning screen outputs. Nevertheless, everyone can benefit from easy to read information that follows a clear and coherent structure. From a certain degree, people with motor disabilities use a computer without keyboard and / or mouse, which means that the use of assistive technologies becomes a sine qua non. Deriving from this fact, menu items, buttons and controls have to be reachable without mouse, thus as a conclusion, particular attention has to be paid to:

- Comprehensive implementation of shortcuts and hotkeys – mandatory,
- Reader support (or even speech output) – beneficial.

2.7 Conclusions – User Needs

Starting from the statement: "*As the needs of different groups of people with disabilities seemed to vary very much and were evaluated as sometimes contradicting, it was necessary to define categories of user groups to be able to address them properly*", we realized after working on detailed user needs, that except for few specialized requirements, the 'special' needs of ALL target groups can be described with the following set of guidelines – quasi the greatest common divisor [5]:

- Output and translations in Easy to Read and subtitling of spoken information and / or reader support,
- Visual output of auditory information (system / error messages, alarms…),
- Visual support (attraction of attention, support of the information displayed…),
- Readability / legibility support,
- Colour schemes and contrast in line with better legibility / readability and multimedia implementation,
- Typographic framework (fonts, font size, font weight, kerning…) particularly suitable to ease the process of reading and comprehension without early fatigue,
- Clear and coherent structures,
- Flexible training settings that allow this target group to learn at their pace.

3 Technical Implementation

Starting point in 2004 was the Bit Media ECDL 4© multimedia e-learning system:

- In German; Designed with Macromedia Authorware 6.5,
- With animations and sounds,
- With interactive elements and tasks to be fulfilled within the tool.

After a testing phase, it became clear that all common requirements we postulated above are, to some extent implemented in the original or realisable over user profiles within the given framework.

For effective Braille display support and support of mainstream speech output we decided to keep the given design and adapt it to an optimized HTML version with the identical look and feel of the original product (enriched by shortcut and hotkey support) for people working with those devices. All other user groups work with an adaptable multimedia version. Please see the following figure for further information on possible adaptations for the single user profiles.

The most apparent difference between multimedia and HTML version is interactive elements in the multimedia product that needed adaptation and redesign.
We rose this challenge with the following adaptations for our HTML prototype:

- The information contained in interactive elements (tips, image maps, tasks, quizzes, glossary) has been added to the main window or linked to a corresponding HTML document,
- Comprehensive shortcut and hotkey information has been added where necessary (for example to reach menu items or special functions),
- Tutorial tasks have to be fulfilled in the original (office) application, task files and sample solutions are attached to the tool,
- Full implementation and usability of glossary information (HTML),
- Access keys to control all important functions.

To fulfil tutorial tasks within the training environment has been evaluated as more convenient and leads to a faster success (in some tools, it is not even possible to click anything but the right 'button' on the screen) but the advantage of fulfilling tasks and tutorials within the original application is that users get used to applications' look and feel and have the possibility to explore it on their own – leading to an in depth knowledge and lasting learning effect.

The possible side effect of feeling 'lost' and 'not accompanied' by an interactive tutorial setting has been cushioned by:

- Adding original files the learners can work with.
  The users do not have to write own lines before practising for example the commands 'copy' 'cut' and 'paste' but open existing sample files also used for the multimedia version to work with.
- Providing sample solutions to check the results.

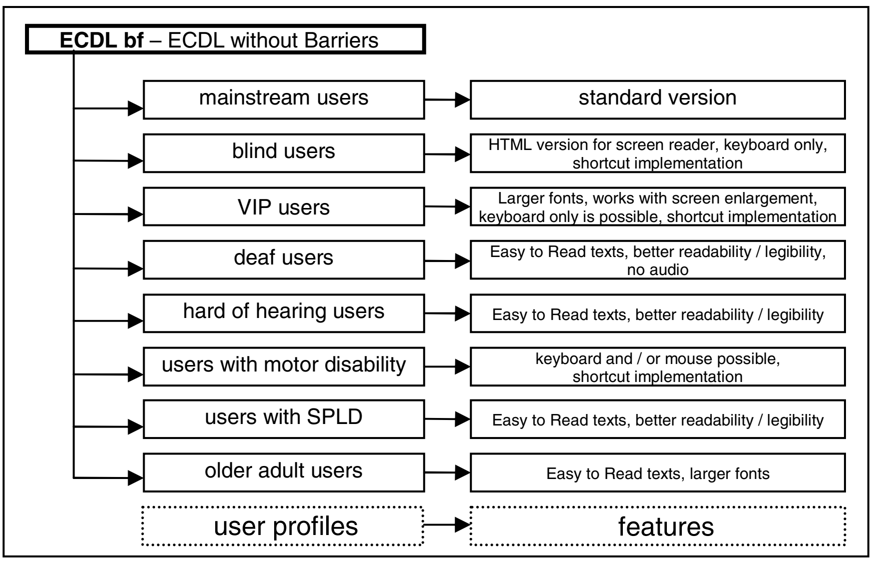


Fig. 1. Possible user profiles and corresponding adaptations to the system

4 Outcomes

The translation and adaptation process (Easy to Read texts and HTML versions with shortcut / hotkey support) has been finished in December 2005. In late 2005, sound recording (multimedia and reader support) and (re)assembly started after user tests on our prototype [6]. The screen (please see figure 2 for details) consists of four major parts: A header part with details (chapter, section, title); a main window with pictures, summarized information, tasks and quizzes; an information window with textual description of the spoken part and additional information / tips; and finally a navigation section.

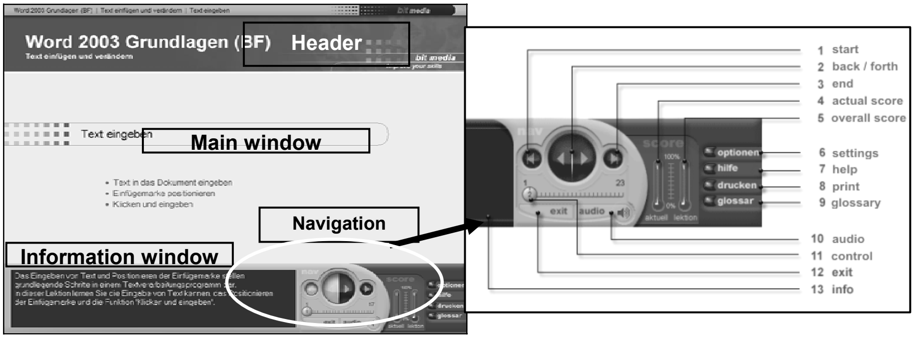


Fig. 2. The prototype and a detail screen of the navigation section (multimedia version)

Access Keys simplify orientation and navigation. The following access keys are provided in the standard version:

Table 1. Access keys and assigned actions

Alt+1	Alt+2	Alt+3	Alt+4	Alt+5	Alt+6	Alt+7	Alt+9
start	back	forth	end	settings	print	glossary	exit

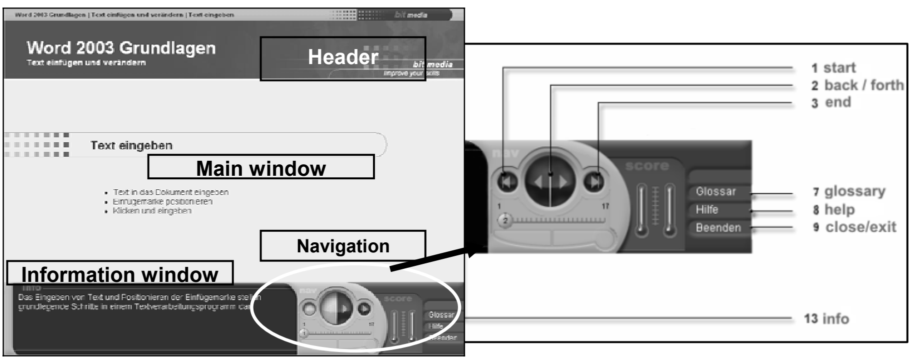


Fig. 3. The prototype and a detail screen of the navigation section (HTML version)

Provided Access keys (HTML version):

Table 2. Access keys and assigned actions

Alt+1	Alt+2	Alt+3	Alt+4	Alt+7	Alt+8	Alt+9
start	back	forth	end	glossary	help	close window

The multimedia version can be adapted to the users' needs. Possible adaptations are Font Size (Standard, Enlarged); Contrast (Contrast 1 and 2); Text version (Standard, Easy) and Gesture video (Active, Inactive). All combinations of adaptations are possible at once.

5 Next Steps

Already during its development, the project 'ECDL bf' attracted considerable interest in Austria and the neighbouring German speaking parts of Europe. The possibility to use one and the same (accessible) teaching and training system (as CD-ROM and possible online version) on the mainstream market as well as in inclusive and / or specialized teaching and training settings has been valued.

In January 2006 intensive negotiations started to transfer 'ECDL bf / ECDL without barriers' to Germany and (primarily the German speaking part and then all parts of) Switzerland.

Italy and France are to follow and already showed strong interest in a co-operation.

In the near future, the consortium heads to adapt the system to English and to disseminate the product throughout the 'ECDL Universe'.

References

1. Internal project paper, May 2003
2. Austrian Report on Disability, 2003
3. Internal project paper, May 2005
4. Miesenberger, K., Petz, A., Leahy, D.: ECDL PD: Access to a Standard Certificate Based on Universal Access to HCI, in: Stephanidis, C. (ed.): Universal Access in HCI – Inclusive Design in the Information Society, Volume 4, Lea, Mahwah, 2003
5. Miesenberger, K., Petz, A., Leahy, D.: ECDL PD: A Standard Certificate for All, in: Craddock, G., M., McCormack, L., P., Reilly, R., B., Knops, H., T., P.: Assistive Technology - Shaping the Future, Proceeding of the AAATE Conference 2003, IOS Press, 09'2003
6. ECDL bf Website with online prototype: `http://www.barrierefrei.ecdl.at`

Searching Knowledge CinemaSense as a Case Study in Collaborative Production of a WWW Service in Two Universities

Antti Raike

University of Art and Design
Hämeentie 135 C, 00560 Helsinki, Finland
http://www.uiah.fi/~araike/
http://elokuvantaju.uiah.fi/

Abstract. In this paper I will present design research carried out between 1999–2004 at the University of Art and Design Helsinki in collaboration with the Classroom Teacher Training Programme for Finnish Sign Language Users of Jyväskylä University, Finland. The aim of the project was to produce an accessible web-based study product, as well as to clarify the sign language students' deepening of knowledge and conceptualization related to the subject of cinematic expression, as well as their collaboration during the web-based course. The aim of the design research was connected to the general aim of inclusion, for a shared university for all, which adapts flexibly to the needs of different and diverse students. The design research was positioned in the areas of film art and pedagogy. By merging participatory action research and WWW production a collaborative study concept dealing with cinematic expression entitled, *CinemaSense*, was developed and produced as part of the research work. It can be accessed at http://elokuvantaju.uiah.fi/. The usability and accessibility of the *CinemaSense* was observed during web-based courses in cinematic expression during 2001, with the help of a concept survey and network-based communication.

1 Design Research and ICT Tools for Accessible CSCL

Collins & al. [1] argue for the need to develop design research and science to investigate how different designs of ICT affect dependent variables e.g. in teaching and learning; just as in aeronautics, where researchers look at how different designs affect dependent variables. Physics, biology, and anthropology can be viewed as analytic sciences, where the effort is to understand how phenomena in the world can be explained. Aeronautics, AI, and acoustics can be viewed as design sciences, where the goal is to determine how designed artefacts behave under different conditions.

Design begins by asking, who is the artefact for and do they need it [2]? *Design for all* (DfA, universal or inclusive design) links to the concept of an *inclusive society*; it is a broad-spectrum solution and a part of modern, multi-cultural society,

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 568–574, 2006.

meant to help every citizen. DfA is an approach to the design of products, services and environments to be as usable and as accessible as possible by everyone regardless of age, ability, culture or situation. *Usability* describes how easily an artefact can be used by any type of user. *Accessibility* describes how easy it is for people to get, use, and understand artefacts. However, in HCI (Human-Computer Interaction) or in ICT (Information and Communication Technology) accessibility may refer to the usability of ICT by people with disabilities. Due to this it is rare to find a DfA product for learning that is used mostly by non-disabled people; products and services marketed as having benefited from a DfA process are often actually the devices customized specifically for use by people with disabilities. Hence, a term adaptive technology could be used more precisely when hardware or software is used to customize a computer for a disabled person [3].

Thus DfA is an inclusive and proactive approach seeking to accommodate diversity in the users and usage contexts of interactive products, applications and services, starting from the design phase. DfA becomes essential as life expectancy has risen and modern medicine has increased the survival rate of the population. The knowledge-intensive information society itself changes radically the way students study and interact with each other and with information [4]. DfA in higher education augment fully able students as well as students with disabilities to access content in their preferred way: e.g. it is rather easy for students with normal vision and with vision impairment or dyslexia to adjust text sizes and colours of an interface [5]. Furthermore, Namatame & al. [6] aimed to design web-based interactive educational materials for the hearing-impaired based on their interaction style. The results of an eye-tracking experiment demonstrate behavioural differences between hearing-impaired and hearing students when using web-based educational materials, which might suggest that the design of web-based materials is insufficient for the hearing-impaired. Another challenge of DfA is to make ICT tools for CSCL (Computer Supported Collaborative Learning) accessible to students with cognitive disabilities; including those with poor communication or reduced reading skills. Similarly there are several types of unobtrusive or disputable cognitive disabilities which are more difficult to diagnose than hearing or vision impairment or colour blindness: dyslexia or autism, motor or dexterity disability such as paralysis, cerebral palsy, and carpal tunnel syndrome. These all have impact on the design and use of ICT in higher education. But what is good Design for All practice in higher education? How do you design a service for all possible students?

2 Research and Methodological Approach

The research problem to solve was how to produce an accessible WWW service on cinematic expression, which supports both collaborative web-based learning as well as individual development of knowledge in the field of film [7–9]. The methodology was based on user-centric and participatory design methods to launch a collaborative design research [10–12]. At the commencement, a service concept for flexible collaborative web-based study of film was developed and the first version of the *CinemaSense* portal was produced [13] (fig. 1). Secondly, the WWW service was tested and developed into an accessible and multi-cultural, art subject, web-based study format with the help of Deaf students [14–16]. Thirdly, observations and experiences from the design research were sought for theoretical consilience with the help of constructive and collaborative learning theory, interdisciplinary with the cognitive sciences and research of cultural evolution. [17]

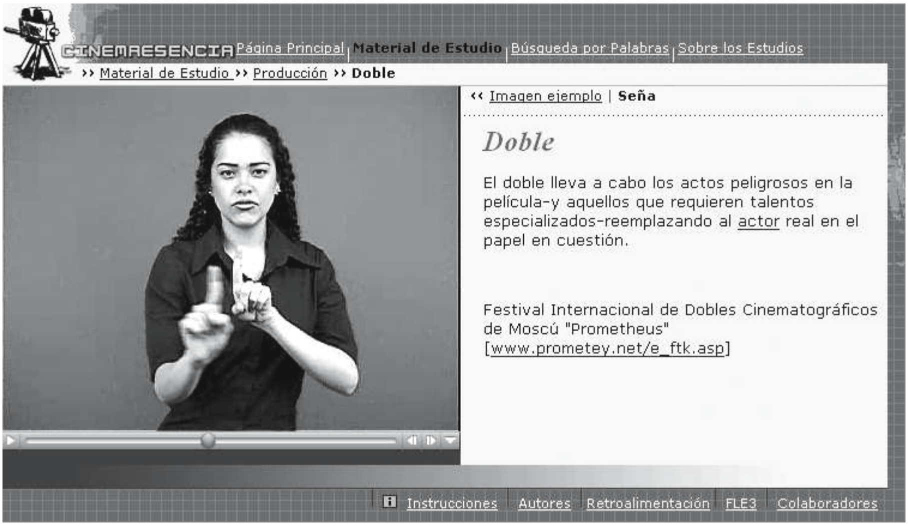


Fig. 1. Spanish version of the CinemaSense – CinemaEsencia – with Colombian Sign language, `http://elokuvantaju.uiah.fi/`

A grounded theory approach was used where a textual database or corpus (such as emails of the participants) was read and re-read to discover or label variables and their interrelationships [18]. The data of the research was made up of the concept maps [19], questionnaires, e-mail messages, diaries and documentary films of the research group participants as well as the concept maps and initial questions of the control group. Beside the corpus, all aspects of the inquiry learning process, i.e. setting up research problems, constructing ones own working theories, searching for new scientific information, was shared with fellow students by using a shared CSCL database during studies. The information objects, e.g. in the form of notes, were produced as a dialogue so that each note was commenting and linking to another. The metadata of each note like, the author's name, the category of inquiry defined by the author, and to which note it is referring to was created during the dialogue. Finally the data was complemented with the e-mail communication of the *CinemaSense* production group members as well as the MA thesis project, dealing with the *CinemaSense* production process. As a result, there was information in the large data corpus of the Deaf students' emails, learning environment (FLE) [20] discussion messages and concept maps drawn by each student during film study workshops.

Different self-organizing maps (SOM) were generated based on the metadata and the written information in the subject and body of the notes [21]. The SOM is a widely used artificial neural network model. Similarity clustering achieved by means of the SOM is suggested as an alternative or complementary method to conventional list-based approaches of organizing large data characterized by significant patterns of multiple criteria [22, 23]. Finding structures in vast multidimensional data sets like students e-mail messages is difficult and time-consuming. Kohonen's SOM can be used to aid the exploration; the structures in the data sets are illustrated on map displays where similar items lie close to each other [24]. The SOM learning process is unsupervised: no a priori classifications for the input examples are needed. The learning process is based on similarity comparisons in a continuous space. The result is a system that associates similar inputs close to each other in the two-dimensional grid called the map (fig. 2).

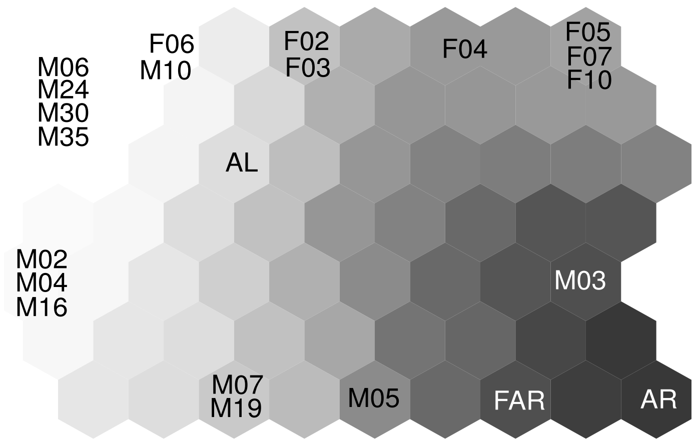


Fig. 2. A map of FLE learning environment and e-mail communication based on strategies of collaboration. F=FLE message, M=e-mail, 01–10=students of Jyväskylä; 16–35=students of Helsinki; AR=Tutor, AL=tutor of editing workshop. Note student 03 using e-mails (M03) near AR and clusters of students on the left of the map. The student 03 collaborates more in the learning environment (F03).

The input may be highly complex multidimensional numerical data [25]. The corpus was coded for the SOM analysis. The maps were then used to evaluate how students conceptualization evolved during the film studies in distance education, an important aspect of a collaborative and complicated project. Hence the SOM was a tool for getting an overview of the students learning and development of conceptual thinking, in this case concerning film art. From the researcher's viewpoint, it was difficult to find coherent patterns in such complicated data, because the significant information had to be collected from a vast corpus of written texts. Hence presenting the students knowledge with maps reduced the cognitive load of the researcher.

3 Findings and Products

1. Theoretical investigation and description of the design research process as to how to produce a web service and basic level, net-based learning material for the study of cinematic expression.
2. The research group participants' concept maps became more conformed and their concepts became more professional whilst studying cinematic expression in the web-based course.
3. By the end of the web-based course the concept maps of the participants had developed from film viewer maps to filmmaker maps.
4. The research group traced professional production, designed and developed the subject of a documentary film, organized the production and produced three documentary films.
5. The accessible and multi-language *CinemaSense* v1.0 service and a web-study course concept for flexible art subject studies was produced. The research also increased understanding in interactivity of multi-lingual, web-based study. The results can be applied in the production of multi-modal web courses, interfaces and services that, for their own part, promote inclusion as well as multi-cultural and flexible university study.

4 Future Plans

Design and research solve the same problems of quality and innovation [9]. However, unlike research, design does not have to be new although it has to be good from the users perspective. A good design is not imposed on the students without first studying the intended users and figuring out what they need. Future design research projects should bring the concept Design for All (Universal design, inclusive design) under critical research by methods of practical design experiment. Thus it would be possible to discover a robust rationalization for the rather political concept of DfA which is mainly derived from the rhetoric of modern welfare state and information society.

The methodology of the design research should be developed as an interdisciplinary effort by reflecting data with the results of e.g. cognitive sciences and brain research. Further research should clarify the relation of so-called cultural universals, varying biological constraints and situated cognition for the collaborative design task. Researchers, designers, professionals, and students should codesign multimodal, multicultural, and multilingual interfaces for advanced studies in higher education. Design tasks should involve designers representing variable cultural backgrounds, and accessibility of the designed artefacts should be tested with users of diverse cultural and physical qualities. Hence design research would contribute to production of accessible and universal ICT tools for CSCL in higher education.

References

1. Collins, A.; Joseph, D. & Bielaczyc, K. (2004). Design Research: Theoretical and Methodological Issues. Journal of the Learning Sciences, 13 (1), 15–42.
2. Díaz-Kommonen, L. (2002). Art, Fact, and Artifact production. Design Research and Multidisciplinary Collaboration. Publication series of the University of Art and Design Helsinki A 37. Helsinki: UIAH.
3. Miesenberger, K.; Klaus, J.; Zagler, W. & al. (eds.) 2004. Computers Helping People with Special Needs: 9th International Conference, ICCHP 2004, Paris, France, July 7–9, 2004. Proceedings. Lecture Notes in Computer Science. Volume 3118/2004. Heidelberg: Springer-Verlag.
4. Hakkarainen, K.; Palonen, T.; Paavola, S. & Lehtinen, E. (2004). Communities of networked expertise: Professional and educational perspectives. Amsterdam: Elsevier.
5. Nielsen, J. (2000). Designing Web Usability. Indianapolis: New Riders
6. Namatame, M.; Kitajima, M.; Nishioka, T. & Fukamauchi, F. (2003). A Prepatory Study for Designing Web-Based Educational Materials for the Hearing-Impaired. In Miesenberger & al. (eds.) Computers Helping People with Special Needs: 9th International Conference, ICCHP 2004. Lecture Notes in Computer Science, 1144–1151.
7. W3C – World Wide Web Consortium (2004). Web Content Accessibility Guidelines 2.0. W3C Working Draft 19 November 2004. http://www.w3.org/TR/WCAG(20/
8. Theofanos, M. F. & Redish J. (2003). Guidelines for Accessible–and Usable–Web Sites: Observing Users Who Work With Screenreaders. Interactions, bimonthly publication of the ACM, The Association for Computing, volume x.6, November–December 2003, 38–51.
   http://www.redish.net/content/papers/InteractionsPaperAuthorsVer.pdf/
9. Brown, A. L. (1992). Design Experiments: Theoretical and Methodological Challenges in Creating Complex Interventions in Classroom Settings. Journal of the Learning Sciences. Vol. 2 No. 2, 141–178.
10. Axelrod, R. (1984). The Evolution of Cooperation. New York: Basic Books.
11. Choo, C. W.; Detlor, B. & Turnbull, D. (2000). Web Work – Information Seeking and Knowledge Work on the World Wide Web. Dordrecht: Kluwer Academic.
12. Wenger, E. (1998). Communities of practice. Learning, meaning, and identity. Cambridge, UK: Cambridge University.
13. Raike, A.; Laitinen, K. & Viikari, T. (2001). Elokuvantaju – *CinemaSense*. http://elokuvantaju.uiah.fi/
14. Ladd, P. (2003). Understanding Deaf Culture: In Search of Deafhood. Clevedon: Multilingual Matters.
15. Collins, A. (1990). Toward a Design Science of Education. CTE Technical Report Issue No. 1, January 1990.
16. Brown, J. S.; Collins, A. & Duguid, P. (1989). Situated Cognition and the Culture of Learning. Educational Researcher, 18, 1, 32–42.
17. Raike, A. (2005). Löytäjät Elokuvantajua rakentamassa – yhteisöllinen WWW-palvelun tuotanto. Publication series of the University of Art and Design Helsinki A 57. Helsinki: UIAH.
18. Honkela, T.; Leinonen, T.; Lonka, K. & Raike, A. (2000). Self-Organizing Maps and Constructive Learning. Proceedings of ICEUT'2000, International Conference on Educational Uses of Communication and Information Technologies, Beijing, China, August 21–25, 2000, 339–343.

19. Murtonen, M. & Merenluoto, K. (2002). Concept maps as a tool for tracing the process of conceptual change – a methodological perspective. Proceedings of Third European Symposium on Conceptual Change. A process approach to conceptual change. June 26–28, Turku, Finland. 56–61.
20. Leinonen, T. & al. (2003). FLE3 – Learning Environment. A web-based groupware for computer supported collaborative learning (CSCL). http://fle3.uiah.fi/
21. Kohonen, T. (2001). Self-Organizing Maps. Springer Series in Information Sciences 30. Third Edition. Berlin: Springer.
22. Kaipainen, M.; Koskenniemi, T.; Kerminen, A.; Raike, A. & Ellonen, A. (2001). Presenting data as similarity clusters instead of lists. In: Stephanidis, Constantine 2001. Universal Access in HCI: Towards an Information Society for All. Mahwah, NJ, London: Lawrence Erlbaum.
23. Kerminen, A.; Raike, A. & Kaipainen, M. (2000). Self-organizing map browser for database retrieval. Consiglio Nazionale delle Riserche, Istituto di Ricerca sulle Onde Elettromagnetiche. Proceedings of 6th ERCIM Workshop, User Interfaces for All. Convitto della Calza, Florence, Italy, Oct 25–26, 2000.
24. Ultsch, A. (1993). Self-organizing Neural Networks for Visualization and Classification. In Opitz, Lausen & Klar (eds.). Information and Classification, Berlin, Springer, 307–313.
25. Barabási, A-L. (2002). Linked: The New Science of Networks. Cambridge, MA: Perseus.

Deaf and Hard of Hearing People: Electronic Communication Aids

Introduction to the Special Thematic Session

Hans-Heinrich Bothe

Technical University of Denmark, Centre for Applied Hearing Research,
Oersteds Plads b.352, 2800 Kgs. Lyngby, Denmark
hbb@oersted.dtu.dk

Abstract. This paper gives an overview and extends the Special Thematic Session (STS) on 'Deaf and Hard-of-Hearing People: Electronic Communication Aids'. The topics of the session focus on special equipment to improve the communication between people with or without hearing-impairments, on improvement of communicational skills, or on methods to generate a better research environment for respective technical developments. The papers are related to visual communication using sign language, lip-reading or written text, to spoken or written language acquisition or improvement, or to general services for hearing-impaired persons.

1 What Does the Word 'Hearing-Impairment' Mean?

About six million people with severe hearing-impairment live in the European Union. Traditionally, we distinguish between hard-of-hearing, late deafened or deaf people. The major difference in terms of language skills is determined by the question whether the impairment occurred before or after basic acoustic language acquisition. The mother tongue for hard-of-hearing or late deafened people is an acoustic language, whereas the one of deaf people is a respective sign language. Since sign languages are visual languages and language skills strongly influence abstract thinking and thus, the way of thinking, communication problems between the two groups are likely to occur. Many hearing-impaired people organize in associations either for deaf or for deafened or hard-of-hearing. The deaf society is traditionally very strong with several special interest groups, whereas hard-of-hearing or deafened people tend to stay in their 'old society', i.e. the people without hearing-impairments.

The consequences of this splitting yield an ill positioning in the education of children in many European countries, where educational dogmas are sometimes imposed against individual help for the children. Teachers are often urged to use signing (sign language or language-related sign language) or not signing (speech reading), dependent on country, region, school, and influenced by their individual university education. Developers of technical aids often stumble 'responsibility-free' into one of the two groups and follow the own interest in image or speech processing or other language-related signal processing or programming techniques.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 575 – 578, 2006.

The balance of technical innovations is therefore often based on personal interests instead of on educational needs of the children. With the improvement of high-tech supportive equipment, some of these old fixed schemes are becoming less important. Especially the large number of cochlea implant operations in babies or young children has an enormous impact on future developments of political decisions and educational paradigms; many children with cochlea implants can now join ordinary schools and will socialize in both, the deaf and the non-hearing-impaired societies. This new situation will naturally lead to new, non-conformal paradigms in the educational systems and in the assimilation between the two societies. Developers of technical equipment should consider such actual and future perspectives, take-over more social responsibility, and participate in the prospective unification process.

2 Communication

Individuals need to communicate with each other. Naturally, we make use of a variety of communication channels or modalities. The stimulation of modalities leads to respective concerted perceptions or emotions. It is difficult to create similar perceptions or emotions by substituting one modality by another one. The learning abilities of humans may substitute for this deficiency.

As a good painter needs to have good basic painting skills, a good communicator needs to have good basic communicational skills; these essentially concern recognition and association – which may be trained by stereotype computer programs or other technical aids. Abstract understanding is individual and must evolve from an individual person-to-person dialogue. The fundamental domain of our developments is therefore the acquisition of basic communicational skills.

3 Electronic Communication Aids for Hearing-Impaired Persons

Different hearing-impairments lead to different suggestions or technical requirements for the substitution of the hearing loss in order to communicate within the majority-environment of the non-hearing-impaired. In the following, some technologies will briefly be mentioned.

3.1 Cochlea Implants

Cochlea implants substitute the mechanical stimulation of the inner hair cells at the basilar membrane in the inner ear by an electrical stimulation. A distinct number of electrodes are lined-up on a plastic tail, which is inserted into the cochlea. The electrodes cover a wide range of acoustic frequency areas in the cochlea. The complete hearing aid consists of two parts, the external microphone with signal processing unit and the implanted electrodes with control unit. The sound signal is received by the microphone and decomposed into a distinct number of sound channels, the outputs of which being fed through an inductive or radio coil to the single electrodes. Typical numbers of electrodes are 32 or 64. The insertion of the plastic tails with the electrodes destroys any residual hearing, which makes the implantation an enormously important decision for the future of the patient.

Implantations are often applied to babies or children, if possible within the first year of life, when the auditory pathway and the auditory cortex still have their maximum plasticity. Successful operations lead to new type of acoustic language acquisition. Few technical training aids have so far been developed to support language acquisition or certain communication deficiencies of people with cochlea implants.

3.2 Hearing-Aids

People with residual hearing abilities may benefit from external hearing aids. In many cases, the demands of reconstructed directional hearing require binaural hearing aids, i.e. amplifier devices in both ears. These amplifiers must be individually adapted in order to compensate for the individual hearing-losses. There is a wide variety of analog or digital hearing aids with very advanced miniature signal processing units.

Traditional hearing aids substitute certain sound frequencies and intensity levels by other ones and are most often used by people with an ordinary language acquisition.

3.3 Tinnitus Masker

A tinnitus is a sound perception that lacks any external acoustic stimulation. It can be caused by deficiencies in the auditory pathway including the cerebral auditory cortex. Permanent tinnitus sensations strongly influence the physical and psychological state of the patient. Tinnitus maskers are acoustic stimulators similar to hearing aids that repetitively send-out sound pulses or short sound sequences in order to mask the perception of the internal permanent sensation. They must be adapted individually and are worn as ear channel plug-ins or behind the ear mainly at night when the sensation is predominant. The noncommercial Tinnitus Masker [1] is an easy-to-use alternative that allows to choose the sounds that mask the tinnitus best and to mix them in order to create real-time masking relief via headphones. WAV or MP3 may also be recorded to create a custom portable masking relief.

3.4 Training Aids

There are diverse technical training-aids available to improve visual communication skills through sign languages, language-related sign languages, finger spelling, or speech reading (lip-reading) or to improve one's speech intelligibility with the help of a visual feedback system. The two general types of training aids are CDs with i) pre-recorded video material or with ii) visual or audio-visual computer animations. Technical training aids require a good didactical program so that the patients can gain the maximum effort to improve their basic communication skills. Technical training-aids may for a short time be used in self-training but should generally be accompanied by the supervision of a human teacher in order to avoid artefactual education.

3.5 Tactile Training Aids

The idea to support the process of speech reading with the help of a tactile assistive aid is motivated by the success of the Tadoma method of speech reading [2]. Tadoma is a natural vibro-tactile method. It is mainly used for the communication with deaf-blind persons, who perceive articulatory movements and actions that occur during the production of speech with the help of their fingers and hand by touching the speakers face and neck.

At present, there are diverse tactile aids available, which support speech reading by providing extra information that is visually not available. Examples are the MiniVib4 [3] or the Mini-Fonator [4] devices, which consist of vibrator arrays or bands attached to the body (e.g., to the chest). Output of a typical research project in the area is the OSCAR hand-held vibro-tactile aid [5], which presents features of the acoustic signal through two vibrators to the user. An improved version is the handhold FLIPPS with five vibro-tactile fingertip stimulators [6].

4 Topics of the Special Thematic Session

The STS is with ten papers mainly focused on technical devices or applications for hearing-impaired persons, followed by technological methods to develop such devices with three papers and by technical training-aids to develop communicational skills with two papers. The STS covers improvement of visual communication using sign language, lip-reading or written text, spoken and written language acquisition, and general services for hearing-impaired persons.

Acknowledgments

I appreciate very much the patience of our little daughter Henny who had to tolerate her father's absence during parts of her fourth birthday.

References

1. `http://www.zdnet.de/downloads/prg/2/1/en10425221-wc.html`
2. Alcorn, S.: The Tadoma Method. Volta Rev. 34 (1932) 195-198
3. MiniVib4: Special Instruments Development, Box 3050, SE-18103 Lidingö, Sweden.
4. Mini-Fonator: Siemens Hearing Instruments Inc., 10 Costitution Ave, PO Box 1397, Piscataway, NJ 08855-1397
5. Spens, K.E., Huss, C., Dahlqvist, M., Agelfors, e.: A Hand-held Two-channel Vibro-tactile Speech Communication Aid for the Deaf: Characteristics and Results. Scandinavian Audiology, Suppl. 47 (1997) 5-13
6. H-H. Bothe, Ali F. Mohamad, Martin R. Jensen: Electro-vibratory fingertip stimulation for improved speechreading. In: Miesenberger, K. *et al* (eds.), Lecture Notes in Computer Science, Springer-Verlag Heidelberg, vol. 3118/2004, Computer Helping People with Special Needs: 9th International Conference, ICCHP 2004, Paris, France, July 7-9th, 2004, pp. 1152-1159

User Evaluation of the SYNFACE Talking Head Telephone

Eva Agelfors[1,⋆], Jonas Beskow[1], Inger Karlsson[1], Jo Kewley[2], Giampiero Salvi[1], and Neil Thomas[2]

[1] Kungl Tekniska Högskolan (KTH), Dept of Speech, Music and Hearing, SE-100 44 Stockholm, Sweden
{agelfors, beskow, inger, giampi}@kth.se
[2] Royal National Institute for Deaf People (RNID), London EC1Y 8SL, UK
neil.thomas@rnid.org.uk

Abstract. The talking-head telephone, Synface, is a lip-reading support for people with hearing-impairment. It has been tested by 49 users with varying degrees of hearing-impaired in UK and Sweden in lab and home environments. Synface was found to give support to the users, especially in perceiving numbers and addresses and an enjoyable way to communicate. A majority deemed Synface to be a useful product.

1 Introduction

Hard of hearing people often rely on lip-reading to follow conversations. This works well in face-to-face situations but over the telephone this visual information is missing, and people are left to rely only on what they can hear. This means that telephone conversations can be difficult and frequently these problems are greater when the person at the other end is a stranger. Many people with hearing difficulties report that they have stopped using the telephone as their hearing impairment has increased.

A solution a this problem could be video telephony. The current technology for video telephony lacks the quality that is necessary for lip reading [1], and sets a number of problems ranging from lack of privacy to the need for special equipment on both ends of the line. The SYNFACE project has developed a prototype system with a synthetic talking face that reproduces the lip movements of a talker only from the acoustic signal [2]. This can be used by the person with impaired hearing to improve understanding, transparent to the person on the other end of the line. The prototype has been developed for Swedish, English and Dutch. This paper describes user evaluations of the prototype system for Swedish and English with hearing impaired users in Sweden (KTH) and UK (RNID), respectively.

2 SYNFACE Talking Head Telephone

The main parts of the Synface system are a phoneme recogniser and a three dimensional artificial face that runs on a standard pc. As the automatic pho-

⋆ Authors in alphabetical order.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 579–586, 2006.

netic recognition and the visual animation software need to run simultaneously on a compact machine, the computational resources are quite limited. A unique phoneme recogniser was developed for Synface [3] that was optimised to deliver phonetic output with very low latency (30 msec). Recognisers for the three languages were trained on speech collected in the ESPRIT SpeechDat project. The three dimensional artificial face animation is based on work originally described in [4]. An articulatory control model [5] is used to animate the movements of the artificial face. This model takes time-stamped phonetic symbols as input and produces articulatory control parameter trajectories to drive the face model. The control model includes coarticulation to account for the influence of neighbouring segments. A special real-time version of the rule-based control model has been developed, that uses a finite time-window of articulatory anticipation. For the prototype Synface system, control models have been adapted to Dutch, English and Swedish.

In the Synface prototype used in the evaluations the incoming telephone speech is tapped of the line, fed into the computer and analysed to produce the face articulation control parameters. The speech is delayed by 200 msec to allow for the delays in the recogniser and the artificial face, time aligned with the face movements and played back to the user in synchrony with the face.

3 SYNFACE User Evaluations

The focus of the evaluations was on gaining users attitudes and opinions towards Synface. The evaluations reported in the present paper were performed at two places: RNID performed the evaluations of the English prototype and KTH of the Swedish prototype. Between the two sites, a total of 49 evaluations have been completed. These can be broken down into lab-based and home based trials, as described below.

The structure of all tests were similar although the duration and locality differed. The evaluation participants were informed about how Synface worked, they performed some comprehension tests and made telephone calls using the Synface prototype. In connection to the tests the evaluation participants were asked to fill in a set of three questionnaires:

- *Selection questionnaire.* Designed to find out basic information about potential participants to ensure that a variety of users were recruited for the evaluations.
- *Pre-Evaluation questionnaire.* Designed to gain an understanding of each participant, how they use existing communication methods and the problems they have encountered.
- *Post-Evaluation questionnaire.* Designed to gain the participants' opinions of Synface, the problems they encountered, the usability of the prototype, whether they would use such a system in the future and additional marketing related questions.

The questionnaires were compiled by RNID in English and translated into Swedish to facilitate comparisons between the evaluation results. In addition to filling in the questionnaires all participants were interviewed after the evaluation sessions.

3.1 British Evaluations

The evaluation participants were selected from a larger group who had volunteered to participate. They had filled in the selection questionnaire and were selected to form a group with an even spread of hearing impairments and a bias towards the older age groups.

Lab-Based Evaluations. Extensive lab based evaluations were carried out with 33 participants (14 males, 19 females). Of these, 25 used hearing aids, 5 had a cochlear implant (CI) and 3 participants used no listening devices. Participants were asked to describe their hearing level (when using hearing aids or CI, if applicable). Fig. 1 illustrates the distribution of participants' responses. To gauge if and how participants use lipreading they were asked a number of questions. 97% of participants stated that seeing a speaker's face helped them to understand conversations. 36% felt they were able to understand without seeing the talker's face if in a quiet room, whereas only 21% felt they would be able to in a noisy room. They also differed in how much they used voice telephones. Five participants no longer used a telephone, 4 participants used the telephone less than once a week and 4 participants used a voice telephone weekly. Only 48% of participants felt confident using a voice telephone. The main reason for not using the telephone was that they could not hear enough. Results also suggested that communication with familiar voices is easier than with strangers: 51.5% of participants felt that they could easily understand people they knew on the telephone whereas only 15.2% felt they could easily understand people they did not know. Several participants stated that they struggled understanding names, numbers and people with accents.

The evaluations were carried out in a research lab at the RNID, and the same experimenter carried out all the evaluations. In each lab evaluation participants

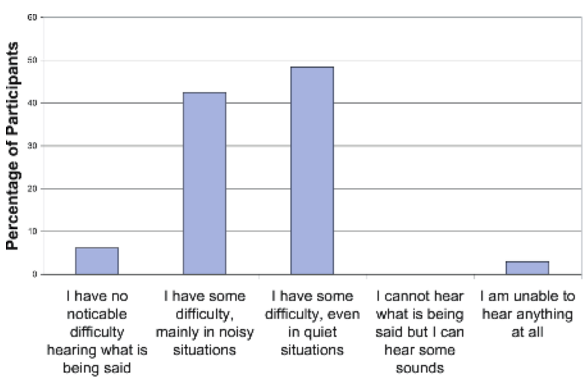


Fig. 1. British participants' descriptions of hearing loss with hearing aid or CI

were first asked to complete the pre-evaluation questionnaire. They were then shown the Synface video and any questions in relation to Synface and the evaluation were answered. Participants were then shown the short story about the North Wind.

The experimenter then described each of the four scenarios that would be the topics of each of the calls they would make using the prototype. Each participant was given the same examples of possible questions they could ask. A second experimenter was at the other end of each of the calls. After participants had completed the scenarios they were asked to complete the Post-Evaluation questionnaire and then a full debriefing interview was carried out.

Home-Based Evaluations. The second part of the UK evaluations was to investigate how participants' attitudes and opinions changed with extended usage and whether users' ability to lip-read the face improved with practice. 6 people, selected from those that participated in the lab-based tests participated in these evaluations. These home-based evaluations involved the prototype being installed in a participant's home for a period of 2 weeks. The participants reported either mild or moderate hearing losses. Four of these participants stated that they find communication using a voice telephone "somewhat effective" and 2 stated that they find the telephone "not at all effective". Of these participants 3 use a voice telephone on a daily basis, 1 uses a voice telephone on a weekly basis, 1 uses a voice telephone less than once a week and 1 participant no longer uses a voice telephone. In addition to questionnaire results, data was collected using an informal interview at the start and end of each evaluation as well as a short telephone interview two weeks after the evaluation had ceased. Quantitative data was also obtained by giving participants a short questionnaire based on the post evaluation questionnaire used in the lab-based evaluations.

To gauge whether any learning effects took place, participants were scored on the number of keywords understood in pre-recorded sentences. These sentences were presented to participants at the start and end of each evaluation.

3.2 Swedish Evaluations

The Swedish evaluations were either lab-based, 5 participants, or home- or work-based, 5 participants. The lab-based evaluations lasted about half a day and contained different tasks including comprehension tests, a telephone call to one experimenter following a set scenario and unsupervised calls. The home- or work-based evaluations ran for one to four days. These evaluations contained some initial familiarisation with Synface and how it worked and a call to the same experimenter following the same scenario as in the lab-based evaluation, otherwise they were unsupervised. The home-based test participants also took part in a comprehension test at KTH after their evaluations.

The evaluation participants were recruited among patients at a cochlear implant clinic and a hearing rehabilitation clinic in Stockholm. All participants reported that they had problems understanding unknown people on the phone while one person from each group reported having problems understanding people they know. Their lipreading skills were estimated in the comprehension test, Fig 2.

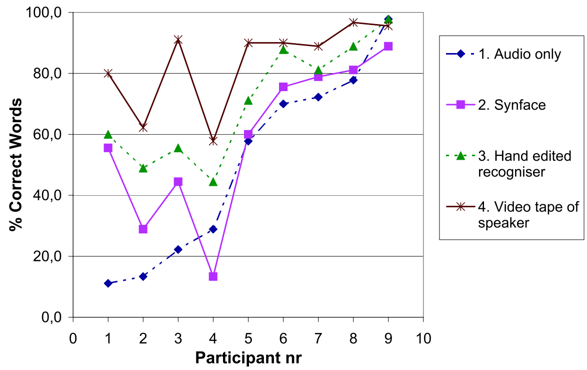


Fig. 2. Word comprehension of 1. speech only, 2. speech with Synface, 3. speech with hand edited talking head synthesis and 4. natural visual speech. All speech is from the same speaker. Results from the Swedish evaluations.

Lab-Based Evaluation. The five subjects, two women and 3 men, that participated in the KTH lab-based evaluations had some prior experience of the Synface technology. Four of the subjects used hearing aids equipped with t-coil for the telephone conversations.The fifth subject had a cochlear implant (CI). Two subjects stated that they could not hear what is being said but could hear some sounds, one that he had some difficulty, even in quiet situations and two had some difficulty, mainly in noisy situations when they were asked to describe their hearing level in the questionnaire. The telephone usage varied between the subjects, one person had stopped using the telephone two years ago, one person used a telephone weekly and the remaining three used it most days.

Home-Based Evaluation. Five evaluation participants, two males and three females, tested Synface in this series, one participant used Synface at work, the other four used it at home. Three of the participants used a CI, the other two used hearing aids. All had a severe to profound hearing loss. One participant stated that she had no difficult in hearing what is being said even though she is using a CI and is relying on lipreading to understand speech. One participant stated that she had some difficulties even in quiet situations and the remaining three had some difficulties hearing speech, especially in noisy situations. Three of the subjects used a telephone every day and two used it weekly. Four of the participants used either hearing aid with a T-switch or a CI when using the phone while the fifth participant used a telephone with amplification.

4 Results

4.1 British Lab-Based and Swedish Evaluations

The evaluation results from the British lab-based tests and the Swedish tests have been combined in the following. Overall the participants found that Synface gave some support. In general attitudes to Synface varied with 17% finding Synface helped a lot, 65% finding Synface helped a little and 19% finding it did not help. Overall only 8% of the participants stated that they felt that Synface did not make communication clearer when using a voice telephone and only 6% stated that they felt Synface did not make communication more effective. When participants were asked before using the prototype if they felt that Synface could help them use the telephone more successfully, responses were all generally positive or neutral. This suggests that if the prototype was improved, there is the potential for Synface to assist many hard of hearing people to use a voice telephone more successfully. This suggestion is also supported by the 80% of participants who used the Synface prototype and stated that it was useful product and the 63% that stated that it was an enjoyable way to communicate. The questionnaires asked participants about problems that they encountered while using the Synface prototype. A large number noticed that Synface made errors, although there was slight variation in the number of errors noticed. The errors that were noticed were found to only upset or annoy 38% of participants but caused confusion for 54% of participants. Comments made after the UK home based evaluations suggest that with time participants may be able to learn to accommodate some of these errors. Each of the evaluations also highlighted that participants showed a preference for a more realistic face and 67% of participants stated that they felt that the movements of the face were not natural. However other qualitative comments made by participants could suggest that these responses were due to the fact that facial expressions indicating for example agreement, turn-taking, mode and emotions, etc. were not available with Synface; that there is a need for greater articulation and differentiation between the visual aspect of the Synface prototype; or that as suggested in the UK evaluations, preference for a more natural or realistic face may be due to personal preference, and not relate to comprehension. When asked if they would prefer Synface to other telephone alternatives, participants indicated that they preferred Synface, and that a videophone was more popular than using a textphone.

4.2 Comprehension Experiments

A comprehension test was included in the Swedish evaluations. In this test comparisons were made between presenting speech as audio only, through Synface or through a talking head with hand edited recognition or as a video recording, Fig.2. The test material has been used in an earlier study [6]. Most users gained some help from Synface compared to no face. Lipreading a real face was found to be more beneficial to comprehension than Synface. This is probably due to the fact that Synface does not display facial expressions to signal agreement, turn-taking etc., which is present with a real face.

4.3 Home-Based British Evaluations

These evaluations have further highlighted the potential of Synface to improve communication for hard or hearing people. The results suggest that some benefit was gained when using Synface. After extended use participants commented that more benefits were gained once they accepted Synface was an aid to listening rather than a replacement, and with time they were able to accommodate some of the errors that were consistently made by the prototype. No learning effects were found to occur over the duration of the evaluations (2 weeks). This may have been due to the length of the evaluation, the measures used or that this skill requires more time to be learnt. However more research would be needed before further conclusions could be drawn.

The simplicity of the prototype allowed participants to interact with the technology with minimal distractions, the importance of which was highlighted by the increased levels of concentration that participants felt were needed using Synface compared to a standard voice telephone. A simple prototype also allowed the investigation of users attitudes towards the concept of a lip readable telephone rather than specific features of the prototype. Although participants showed a preference for a more realistic face during the lab based evaluations, this was seen to be less important after extended use. However a more realistic or natural looking face would be more enjoyable to look at.

After using Synface for two weeks, results highlight the importance of the lip movements being as accurate as possible and that animations are in sync with the audio. Comments were also made that these lip movements need to be emphasised so that it is easy to distinguish between different sounds. The switching mechanism that is used to prevent the delayed audio being played back to participants was also noticed more frequently with extended use and was still seen as not being ideal by the majority of participants.

Despite a number of participants who commented that they made more calls than they would normally whilst using Synface, they felt that once Synface was removed they would quickly slip back into their old ways. Synface did effect how participants communicated, and that with a number of improvements lipreadable devices such as Synface may be desirable and valuable pieces of technology, improving the chances for hard of hearing people to have equal access to voice communication.

5 Conclusions and Future Work

From the results of the evaluations it is difficult to indicate exactly who found Synface the most or least beneficial. Lipreading ability was not found to relate to those who benefited from Synface. However, the UK lab and the Swedish results appear to suggest that those participants, who are able to use a standard voice telephone successfully, gained less benefit from Synface than those who encounter greater problems. The UK evaluations also suggest that those participants who struggle to use a voice telephone still struggle when using Synface. A number of participants, including those who gained very little benefit from using Synface in each of the countries, also stated that Synface was useful with specific details such as names and numbers that they normally struggle with.

The Synface prototype is currently undergoing usability improvements, and a pre-release version of Synface for IP-telephony will be offered to hard of hearing users in Sweden and UK during spring 2006, by newly founded company SynFace AB `http://www.synface.com`.

Acknowledgements

The SYNFACE project is financed by the European Union (EU) under the FP5 IST Key Action I Programme: Systems and Services for the Citizen. The project also benefits from an equipment donation within the HP VoiceWeb Initiative.

In addition to the authors the following people have been active in the SYNFACE project: B. Granström, K-E Spens at KTH, A. Faulkner, C. Siciliano, G. Williams at UCL, UK, J. Kardach, E. Landström, J-I Lindström at Acapela Sweden AB, Sweden, M. Sheard at RNID, UK, and B. Elsendoorn, E. Rikken, N. van Son, A vWijk at Viataal, the Netherlands.

References

1. Lönnborg, T.: Mobil videotelefoni som kommunikationshjälpmedel för individer med hörselnedsättning. Master's thesis, Linköpings universitet (2005)
2. Beskow, J., Karlsson, I., Kewley, J., Salvi, G.: SYNFACE - a talking head telephone for the hearing-impaired. In Miesenberger, K., Klaus, J., Zagler, W., Burger, D., eds.: Computers Helping People with Special Needs, Springer (2004) 1178–1186
3. Salvi, G.: Truncation error and dynamics in very low latency phonetic recognition. In: ISCA workshop on Non-linear Speech Processing. (2003)
4. Beskow, J.: Animation of talking agents. In: Proc. of AVSP'97, Rhodes (1997)
5. Beskow, J.: Trainable articulatory control models for visual speech synthesis. Journal of Speech Technology **7**(4) (2004) 335–349
6. Siciliano, C., Williams, G., Beskow, J., Faulkner, A.: Evaluation of a multilingual synthetic talking face as a communication aid for the hearing impaired. In: Proc. of ICPhS. (2003) 131–134

Towards a Service Integration Portal for Deaf People

Christophe Ponsard[1], Christiane Broekman[2], Cécile Lamy[3], and Martine Fraiture[4]

[1] CETIC Research Center, Charleroi, Belgium
cp@cetic.be
[2] ISTI, High School for translators and interpreters
[3] CRETH, University of Namur, Belgium
cla@psy.fundp.ac.be
[4] Deaf Federation, Belgium
martine.fraiture@ffsb.be

Abstract. High speed connections have enabled video-based communication and given deaf and hard-of-hearing people the opportunity to rely on the visual modality to better communicate with each other and with hearing people. However a number of obstacles still have to be removed to make this technology really accessible for all: technical expertise for operating a computer, lack of compatibility, poorly integrated services,... This papers describes an ongoing project for providing a better accessibility by (1) supporting multiple kind of user terminals (such as PC, videophone, TV, mobile phone) in an interoperable way; (2) with a simple and user friendly interface, (3) giving access to a number of related services such as directories, video message box, relay centre and signed news.

1 Introduction

Hearing impairment is an invisible disability which is quite frequent: it is estimated that 8% of the population suffer from hearing loss. Deaf and hard-of-hearing people have very specific communication needs which are crucial for their personal development and their successful integration:

- *communication with hearing people*: for signing deaf people, this is a major problem as the spoken/written language of the country is not their natural language. They often have a poor level of written language, cannot speak nor read on lips. For optimal comprehension, it is better for them to rely on a sign language interpreter. However many countries lack of such interpreters and booking one generally requires 2 weeks notice. The availability is even worse in large countries where interpreters spend most of their time in travel.
- *remote communication*: there is no real "deaf" equivalent to the phone. Most solutions (TTY, SMS, email, chat) are text-based and many lack real-time interactivity.
- *access to information*: availability of deaf-specific subtitling and/or sign language interpreting of television broadcasts or video material is still marginal in many countries.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 587–594, 2006.

Videophone technology has been available for a long time but before the generalization of low cost high speed connections, only big companies could afford the cost of multiple ISDN lines and the expensive hardware required. With the rise of high speed Internet, progress in compression techniques and development of low cost web cameras, video communication is now within the reach of every computer user. However, a computer-based solution does not fit the need of every deaf person: less educated or older people will have trouble using such a complex interface.

Considering that the primary interest of deaf people is not to process text or graphics but to communicate and access to information, other kind of terminals such as a phone with a LCD screen or a TV set-top box may provide a simpler interface. Moreover it could also give a unified and device independent access to the set of services deaf are looking for and for which they currently have to rely on a number of different devices or applications (SMS, messenger software, web browser...).

The purpose of this paper is to report on our experience of designing such a portal. Before detailing our architecture and prototype implementation in section 3, section 2 will give an overview of the various services and related technologies, partial similar attempts will also be reported. Finally, section 4 will discuss some perspectives and draw some conclusions so far.

2 Review of Solutions Meeting Deaf Needs

This section is an updated summary of a feasibility study we carried out for the Walloon Ministry of Social Affairs and Heath [10]. It also relies on a large European workshop on this topic [11].

2.1 Video-Based Communication

The main need to address is to enable video communication. The related requirements for signing deaf and hard-of-hearing people are well know and were described before in [2,3], and [4].

- *resolution:* CIF (352x288) resolution is required to catch details such as fingers and face expressions (for sign language) and lips (for lips reading).
- *framerate:* 20 FPS (frames per second) is recommended for a natural conversation. Deaf people can cope with 15 FPS but have to slow down.
- *transmission delay:* experience show that a delay less than 0.5 second is not really perceived. Above 1 second, it becomes a barrier to natural conversation.
- *audio:* audio should be available and synchronized for use by people with hearing remains. An appropriate output should be provided to plug hearing aids.
- *screen size:* 6" screen available on videophones is too small and only allows one person to sign. Standard computer displays (17" and above) are appreciated in full screen mode. TV output is also a good option, even analogue CRT: indeed TV resolution roughly matches the minimal video resolution required.

– *light:* strong light, preferably indirect, is required for the optimal camera operation.
– *centring:* signing requires to focus on the upper body while lips reading requires to zoom on the face. It is useful to have a local view to control its position. Many webcams now come with face tracking: this feature is not adapted for signing.

Those are quite demanding requirements which translates into large bandwidth: about 400 kbps (kilobits per second) for the older compression algorithms (h.263-based). As a new generation of codecs (h.264, mpeg4) is now being deployed and will reduce the bandwidth required. Another major issue is quality of service (QoS): Internet transmission does not enforce QoS unless explicitly supported by the provider and only within its infrastructure. To cope with this, extra bandwidth is generally required. In most countries now support 512 kbps on the slower upstream link and can support video meeting the above requirements. Regarding the user interface, the following terminals are available:

– *computer with webcam:* a recent computer can be turned into a videophone device using cheap webcam. Client software include video enabled messengers or softphone clients (like eConf [13], Intellivic [1]). The later provide a better interoperability (H.323 or SIP). Most are commercial, although a open-source project exists on Linux: Ekiga (previously named GnomeMeeting) [12].
– *TV with video set-top box:* solutions are developed by some manufacturers like Leadtek and D-Link. They are much appreciated as they enable a very convivial use on the TV set in the living room and are simple to use.
– *videophone:* it provides a simple and familiar phone interface with only some extra buttons. The main concern here is the size of the screen.
– *mobile:* 3G phones now support video at a poor resolution but with the fantastic advantage of mobility. To cope with this, deaf are developing a "short sign language" which is roughly the equivalent of SMS and for the same kind of usage [7].

2.2 Text-Based Communication

Text-based communication is currently largely supported: TTY, email, FAX, SMS, computer-based chat. Although imperfect, it should remain a minimal service in all circumstances. Some applications like chat and SMS are quite popular among the youth and can also help practicing written language, provided some efforts are made to use a correct language.

2.3 Sign Language Interpretation

Booking a sign language interpreters can be done by phone (through a hearing person), email, fax or SMS. Videophones open new perspectives [6]:

– *video-relay:* the video equivalent of text relay, allowing signing deaf to carry out a phone conversation with a hearing person through the interpreter.

– *remote interpretation:* a signing deaf in trouble with a hearing person calls the interpreter which carries out the interpretation remotely.
– *booking an interpreter:* can also occur through videophone.

Such services are successfully run in number of countries, mostly in the North of Europe, UK, USA and Australia. Of course, they do not replace traditional physical interpretation: video-relay or remote interpretation are not recommended in all circumstances. Nevertheless this contributes to improve interpreter availability: as they spend less time in travel, they can devote more time to interpreting. Another interest is the ability to work from home using a call dispatcher.

2.4 News

Accessible news for deaf and hard-of-hearing people are available on public television as part of their service obligations. However they generally do not both subtitle and sign the news. In Belgium, the news on French speaking public television is signed and the news of the Flemish public television is subtitled. Subtitling requires more effort but has also a larger audience; signing address a more specific audience with less access to information. News for kids is also signed and is quite popular among the whole deaf community. While broadcasting subtitle is supported through teletext, signed material has to be broadcasted on a full channel, generally not the main channel. In Belgium, conflict with sport events frequently results in cuts, deferments or even cancellations.

A new perspective here is to use make information available online. In UK the University of Bristol as setup such a service on their Deaf Station portal [8]. In France, a similar service is provided by WebSourd [14]. The former relies on Quicktime and the later on Flash technology (easier to install). Other streaming technologies such as RealMedia or Windows Media can also be considered here but all of them are require a computer to view them.

2.5 Deaf Directory

In many countries deaf associations have set up a specific directory enabling deaf people to get in touch with each other by the most appropriate communication media among SMS, FAX, GSM-FAX, email, videophone... This is now generally hosted on a web site with some research capabilities. An example is "Telecontact" (`http://www.telecontact.be`) in Belgium.

2.6 Other Services

Other interesting services can also be considered: videoconference, sign languages dictionaries, forums, e-learning support, emergency service. Due to a lack of place, those will not be detailed here.

3 Towards Service Integration

3.1 Current Lack of Service Integration

In the previous section, we highlighted a number of issues such as having to rely on a computer or interoperability. Looking at the big picture, the major concern

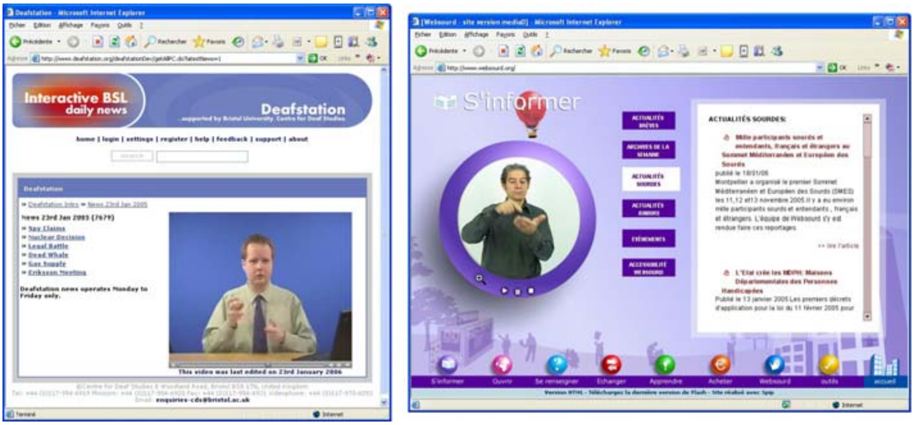


Fig. 1. Deaf Station (left) and WebSourd (right)

is the lack of integration of FAX, SMS, web browser, messenger software. This adds another layer of complexity. Usability could be greatly enhanced through a simple to use interface providing a central access to all those services. Such a portal can be designed to be independent of the user interface and allow the user to choose its favourite terminal (e.g. TV for open conversation, videophone for more private conversation, 3G mobile for use on the move).

WebSourd, already mentioned earlier, has similar goals. It aims at providing a wide range of services structured around information, communication, learning, culture and shopping. However it only relies on a computer-based interface. For now, it is mainly an information portal: interactive services on video such as remote interpreting and conferences are not supported. The portal has a nice Flash-based interface and is bilingual French/French Sign Language which is an important requirement.

In this section, we will elaborated on the design of such a portal and report on the currently implemented prototype.

3.2 Design Issues

The platform architecture is described in figure 2. The portal provides a unified access to various services and content providers in a platform independent way. It enables multimodal (audio, video and text-based) between deaf people and with hearing people.

At infrastructure level, the portal relies on an external SIP-based video over IP infrastructure. SIP is the emerging signalling standard which seems to take over the older H.323 standard [5]. PC interoperability is ensured by the Intellivic [1] softphone which can be embedded within a web browser. Off-line material is also managed by the streaming platform and can be stored in a standard streaming format or in another format and being transcoded on demand.

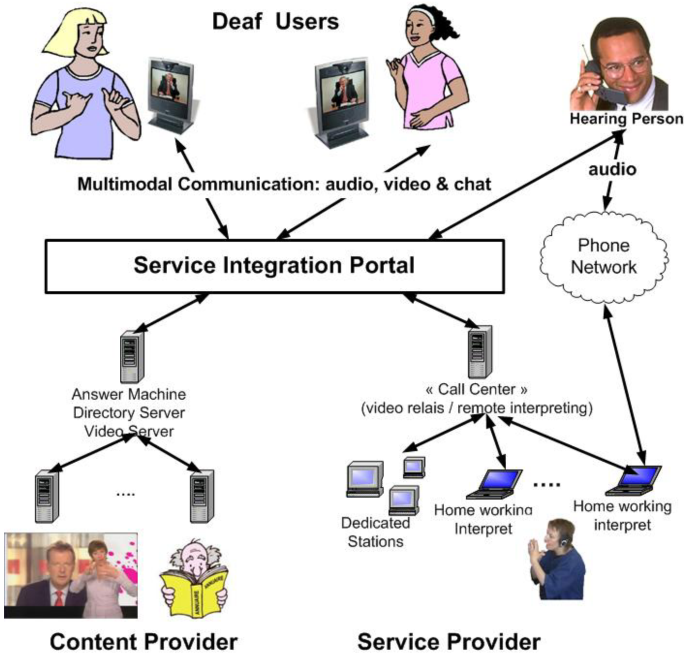


Fig. 2. System Architecture

At user interface level, hardware terminals (videophone, set-top box) have less processing power than a computer. To support the interactivity part of the portal on the hardware, two design choices were considered:

- *server side interface.* The interface is generated from the server. The client is used only to display video streams. However, this requires a toolkit for building the interface (buttons, text fields, etc.) It also places more load on the server and requires to capture user input through some return channel.
- *client side interface.* The interface is managed by the client, typically using an embedded web browser. This requires a more powerful client but the interface can be described using standard HTML. It also provides direct browsing capabilities.

The second design is more interesting and seems now practical: recent hardware is powerful enough and often run some kind of Linux OS with the capacity to support web browsers. It is also closer to the computer-based design on which the portal will be deployed first.

3.3 Portal Prototype

The purpose of this prototype was to experiment with the design and to get feedback from the deaf community and partners who can support the project. The main focus here was on the user interface.

User Interface. The portal is hierarchically organized. Only a small number of information elements are displayed at a time (see figure 3). The navigation zone on the left gives access to video call, directory, messages, access to interpreter, information and emergency call. All the information displayed in French is also explained in sign language in a small video area in the bottom left. The large part on the right of the screen is dedicated to the content: web page browsing or video content (live calls, message or news). This results in a simple and intuitive interface which is also easy to adapt to embedded devices with lower screen resolutions.

Fig. 3. Prototype

Platform Support. The current prototype only fully supports the PC platform as hardware platforms are not yet browser-enabled. However, it is already possible to place video call to/from other SIP-compliant terminals. Video calls are carried out using standard H.263 codec. For now, streaming material is stored in WMV format and not handled by the underlying SIP platform.

The prototype is available on the following web site: http://deafportal.accessible-it.org and requires Microsoft Internet Explorer.

4 Perspectives and Conclusions

The portal was presented to the Belgian deaf federation and demonstrated during events such as the Deaf National Deaf Day and the Handicap & Media workshop [9]. The feedback so far was very positive and we are now proceeding to the next step with the following approach:

- *mature technology:* rely on stable, interoperable, powerful enough devices with an affordable price for the user.
- *progressive deployment:* starting from the communication infrastructure, services will be progressively added based on the technical and organizational issues to solve. The relay centre is currently being designed.

- *stakeholder involvement:* deaf people, interpreters, TV broadcasters, telecom operators are all implied in the development process. For example, the deaf federation is starting a project called Can@lSourd in order to produce news for and by the deaf community.
- *authority involvement:* raising awareness w.r.t. emerging solutions and exploring funding solutions, including social funds for disabled people.

The system will contribute enhance deaf integration and may also produce interesting side effects: new way to work for interpreters and development of a new sign language register. In a wider perspective, the solution can be transposed to home care especially for aging people with services such as health monitoring, emergency call and keeping in touch with relatives.

Acknowledgement

This work was triggered by the Walloon Ministry of Health and Social Matters which financially supported the initial feasibility study. The work was later supported by a Belgacom pilot project. We thanks Androme for giving access to their softphone technology. We also thanks the Belgian French Federation of Deaf People, the many deaf associations (EPEE, Surdimobil,...) and sign language interpreters who have contributed to this work.

References

1. Androme, *Intellivic*, `http://www.intellivic.org`.
2. Frank Blades and Joe Collins, *Remote sign language interpreting, d223, signworks project*, June 1998.
3. Frank Blades and Joe Collins, *Transmission and reception of bsl by videophone, d221, signworks project*, June 1998.
4. Hellstrom Gunnar, *Quality measurement on video communication for sign language*, Proc. of the 16th Int. Symp. on Human Factors in Telecom., May 1997.
5. Alan B. Johnston, *Sip - understanding the session initiation protocol (second edition)*, Artech House, 2004.
6. Thomas F. McCaul, *Video-based telecommunications technology and the deaf community (research summary of the report)*, 1999.
7. Thor Nielsen, *Adoption of Videophony by Deaf People in Sweden*, Information meeting at FFSB, April 2005.
8. Center of Deaf Studies, *Deaf Station*, `http://www.deafstation.org`.
9. C. Ponsard, *Deaf access to media*, Handicap and Media Workshop, December 2005.
10. C. Ponsard and J.F. Molderez, *How can NTIC help Deaf People ? (in French)*, Wallon Region, Ministry of Social Affairs and Health, 2003.
11. Promise European Program, *Preparing a european deaf network for information and communication*, 2000.
12. Damien Sandras and al., *Ekiga*, `http://www.ekiga.org`.
13. France Telecom, *eConf*, `http://www.econf-communication.com`.
14. WebSourd, *WebSourd*, `http://www.websourd.org`.

A Proposal of the Universal Sign Code

Kazuyuki Kanda[1], Tsutomu Kimura[2], and Daisuke Hara[3]

[1] Faculty of Liberal Arts, Chukyo University,
101 Yagoto Honcho,Showaku, Nagoya 466-8666 Japan
kanda@lets.chukyo-u.ac.jp
[2] Toyota National College of Technology
kim@toyota-ct.ac.jp
[3] College of Nursing, Aichi Medical University
daisuke@aichi-med-u.ac.jp

Abstract. This is a proposal of the universal coding system for the sign languages. It is expected to resolve the problems of the entangled notational systems for the sign language. The structure and rules are proposed and some application models for Japanese Sign Language are shown The USC is constructed with several strata, lexical, sentence, phonological, morphological, regional and other levels.

1 Introduction

There are several notational systems for Japanese Sign Language in Japan. The proposed main methods are the following three.

1. Kanda and Atari 1991
2. Nagashima 1998
3. sIGNDEX V.1. & V.2. 2000

Otherwise, internationally, Stokoe's Notation in the USA and HamNoSys in Germany are famous but each country has its own notational system for the sign language, such as in UK and in Sweden.

There is no universal notational system at present, though each system appeals it. As far as comparing those notational systems, there is very little possibility to establish the universal notational system because each country has its own sign language that has its own phonematic environment. As for the vocal languages, IPA is the universal phonetic notational system, not a phonemic notational system. It would take more time to set the IPA of the sign language as long as each sign notational system sticks on describing each sign language.

Moreover, most of the sign notational systems have very complicated and troublesome symbols, which make us to learn a lot and to waste a time. Some of those symbols are for hand writing and not for keyboard use. Therefore they are inconvenient for computer use.

2 The Idea of USC

An idea of the universal code for sign language use (USC below) comes from our effort to learn to utilize those systems and from our dissatisfaction and disappointment. It is

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 595 – 598, 2006.

not our purpose here to propose neither a new notational system nor its revision. We have changed our mind to propose a way to unite them and a way to use them with or without learning the rules.

The USC has several stratums each of which is descriptive to various levels of the sign language; therefore it can be applicable to any sign notational system. In our proposal, the USC has the following stratum of coding levels.

Lexical Level. The codes in this level correspond to a word or a sign. When a user of the USC would like to use the codes simply for the word list or a simple dictionary of a sign language, this level would help. The phonemic or the morphemic level is not easy for a non-expert of linguistic description to handle them. The codes for the fundamental vocabulary would help to compare the signs in the different countries.

Sentence Level. This level is for the sentences of daily use or the basic patterns. The code of the sentence which has the same meaning in other language would help to make a booklet for the travelers, for example. Some certain code for "hello", for example, corresponds to the expressions of Japanese, American, Russian or any sign language.

Phonemic Level. The codes in this level correspond to the phonemes of a sign language, such as the hand-shapes(HS), the palm orientations(PO), the signing locations(LO), the signing movements(MO) and other elements if needed which were widely accepted in the field of sign linguistics. It would help, for example, generating the sign movie connected to the dictionary and the movie generation system through the code. The USC has an open coding system and it is expandable.

Morphemic Level. The codes in this level correspond to the morphemic description the sign language. When the morphemic dictionary would be completed, the morpho-phonemic dictionary is easily created by uniting the phonemic codes and the morphemic code in the morphemic dictionary through the USC.

Regional Codes. Moreover the regional code refers to each level. That enables to the dialects and regional variety.

We propose the concept of the USC and are developing a sample codes as shown in the following.

3 Basic Concept of the USC

At present we propose 13 digits for the USC. We propose to use numerals at present for the codes except regional code, but when expansion is needed, 26 alphabets will be used. Other code area could be attached if a user needs.

The first four digits are for the lexical level that is 9,999 lexical entries. The most of the sign dictionaries in the world contains 2,000 to 6,000 lexical entries. Considering the variety of the meaning to each country, one thousand is regarded enough number.

The following second two digits are for the sentence level that is 99 codes. Our purpose is to select basic expressions of daily use which are quite common in the world. We estimate the basic expressions about 100 sentences.

The third area made of three digits is allotted to the phonemic code. The first of them is the category such as hand-shape, location and so on, and the following two digits are for each phoneme.

The fourth area is for the morphemic codes. In our analysis of JSL, we counted almost five hundred morphemes.

The last area at present is an area code. The abbreviations of each country are widely used now. In the field of internet, two alphabets stand for the country in the URL. We can apply those abbreviations to the USC.

The structure of the temporary USC is as shown in Table 1.

We will show how the thirteen digits are applied. The device we produced for the deaf customer uses the barcode of the USC, which will be demonstrated at the poster session.

Table 1. Structure of USC

Stratum and Construction of USC	
Area 1	Lexical entries
Area 2	Sentences
Area 3	Phonemes
Area 4	Morphemes
Area 5	Region
Area n	Expansion

We can use this USC as a database code of a sign language in many ways. The database would be useful to create or revise a computer aided translation system between sign language and vocal language.

4 The Revised Model of the USC

The above model is a basic idea of the USC which is simple and easy to handle. However we found it is convenient to give a hierarchy to the USC in some case.

a. The USC anticipates QR coding. Not limited to the barcode.
b. A single code is consisted of a length-variable single line.

According to the revised rule above, we propose the following model.

1. The first digit indicates version number.
2. The second two digits indicate the level as the lexical, the phonemic, etc.
3. The third two digits indicate the regional code.
4. The fourth two digits indicate the type of the coding system.
5. The fifth area indicates the information of the descriptive codes.

The revised rules enable the ad hoc or free design of the user of the USC, e.g.

1 ph jp k1 001 2 H1 l B

Here we mean by "001" for the code of "lexis: greeting(aISATSU)", "2H1" indicates the index fingers of the both hands and "l" indicates "the palm orientation is lateral, the final "B" shows the movement of bending.

5 Conclusion

We proposed the basic idea of the Universal Sign Code and its revised model through the process of which we develop the apparatus. We reported some examples of the applications. We also demonstrate the samples of the rules of decoding rules when the USC is applied to Japanese Sign Language. The detail will be presented at the poster session.We would like to continue the completion and revision of the coding system.

Acknowledgement

This paper is supported by the Grant of The Research by the Ministry of Education and Science of Japan titled *Research of the Social Needs of the Deaf and the Proposal from the Cognitive Viewpoint*, in the years 2005-2007, #16091212.

References

1. Kazuyuki Kanda, et.al. "Signdex V.1 & V.2, The Computerized Dictionary of Japanese Sign Language: the notational system and the images of JSL", 7th International Conference on Theoretical Issues in Sign Language Research,, Amsterdam, The Netherlands, (2000) p.65
2. Kazuyuki Kanda, et.al., Notation System and Statistical Analysis of NMS in JSL (In Gesture and Sign Language in Human-Computer Interaction) Lecture Notes in Artificial Intelligence, Springer-Verlag 2298 (2002) 181 – 192
3. Kazuyuki Kanda, Tsutomu Kimura and Daisuke Hara, JSL Typological Glossary and the Appraisal Standard for Sign Engineering Systems, Correspondences On Human Interface, Vol7, No.5, (2005) pp.19-22
4. Kazuyuki Kanda, Tsutomu Kimura and Daisuke Hara, Typological Classification of JSL Signs in the Corpus, IEICE Report, WIT2005-73(2006) pp.79-84

Bilingual Sign Language Dictionary

José L. Fuertes[1], Ángel L. González[1], Gonzalo Mariscal[2], and Carlos Ruiz[2]

[1] Facultad de Informática
[2] CETTICO, Facultad de Informática
UPM, Campus de Montegancedo, 28660 Boadilla del Monte, Madrid, Spain
{jfuertes, agonzalez}@fi.upm.es,
{gmariscal, cruiz}@cettico.fi.upm.es

Abstract. The Spanish Sign Language Dictionary (DILSE) is one of the first truly bilingual (Spanish Sign Language-Spanish) electronic dictionaries for the deaf community. The properties of this format are perfectly matched to a visual language such as sign language, which uses space as a means of expression. Additionally, two-way searches for word entries are possible from either Spanish or signs. The signs have been previously classified according to sign language-based linguistic criteria. Furthermore, the system presented here includes different geographical varieties of Spanish Sign Language.[1]

1 Background

Official recognition of sign language is important for the deaf community. A lot of work has gone into encouraging the dissemination, research and teaching of Spanish Sign Language in particular. Lack of knowledge of this language can lead, on the one hand, to it having a low social status and, on the other, to deaf people not receiving the proper information and training, resulting in functional isolation and thereby preventing this community's full integration into society.

The main goal of the Spanish Sign Language-Spanish Dictionary (DILSE: *DIccionario de Lengua de Signos Española*) is to help people with impaired hearing cope with this adversity. Additionally, it is an aid for anyone wanting to learn or improve their knowledge of Spanish Sign Language (LSE: *Lengua de Signos Española*). This dictionary works like any other bilingual dictionary in that the user can search words by LSE or by Spanish.

To achieve the proposed objective, two clearly separate lines of research had to be set up. The goal of the first was to identify and classify the signs that were to be included in the dictionary. This then raised the need to establish a mechanism by means of which to catalogue and, therefore, be able to search by signs on the basis of their properties. The sign searching method had to be eminently visual, because LSE is first and foremost a visual language.

The second line investigated how to develop software that could be used to create the dictionary, the essential requirement of which was that it would allow searching by sign based on mostly graphical features, the most natural form of expression of which is by means of pictures.

[1] The Spanish National Confederation of Deaf People (CNSE), the ONCE Foundation, the Institute of Migrations and Social Services (IMSERSO) and the Centre of Computing and Communications Technology Transfer (CETTICO) have participated in this R&D project.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 599 – 606, 2006.

2 Related Work

One of the first activities undertaken by deaf communities is the compilation of a sign language dictionary [1]. Limited printed editions of modest-sized basic dictionaries were originally developed. Now, however, thanks to Information and Communication Technology, dictionaries can be developed that take advantage of the whole range of multimedia capability. Representation in print is a weakness, because you are trying to represent statically something (sign language) that is essentially dynamic, and making them bilingual is a tricky point. Additionally, the use of technology not only means taking advantage of the opportunities offered by multimedia, but also making the best use of a wide range of possibilities that databases offer, such as support for two-way searching (LSE ↔ Spanish) and low-cost storage capacity. A number of types of sign dictionaries are defined according to their technological evolution in [1], beginning with printed and ending with multimedia dictionaries, which is the current development trend.

One example of the use of printed dictionaries to create multimedia tools can be found in Australia, where we have "Signs of Australia" [2]. This is an Australian Sign Language (Auslan) dictionary based on the printed edition [3]. It contains about 4000 signs recorded on video and includes sections on the history of Auslan, an introduction to its grammar, and synonyms and antonyms. The dictionary can be searched by English words, by topic or by a subset of phonological properties (hand shape + hand arrangement + location).

The National Deaf Children's Society in the United Kingdom is a clear example of what efforts are being made at standardising the use of sign languages [4]. This organisation has a lot of multimedia tools for use by children and young people.

The creation of a multimedia fingerspelling dictionary for Slovenian Sign Language has led to a standardisation of the use of this sign language [5].

The ASL (American Sign Language) online dictionary is unquestionably a very interesting initiative. It includes over 2500 signs, a figure that grows weekly [6]. Additionally, it is a meeting point for people who want to learn and communicate using ASL, as it includes lessons, a kids zone, ASL-based literature, information directories, travel guides, etc. The online dictionary for British Sign Language [7] is a similar scheme, and now includes around 500 signs.

As regards Spanish Sign Language (LSE) dictionaries, there are a number of editions of printed dictionaries [8], [9]. Although they amounted to a big step forward at the time, they are flawed because they are paper editions.

The "Signos 97-98" dictionary [10] is a LSE dictionary designed to support a sign language learning course. It contains grammatical and usage information and takes an educational approach. It includes about 1000 terms rendered with all multimedia characteristics, and limited bilingual searching is enabled. To search for a Spanish term by LSE, the system is based on the analysis of formative parameters (hand shapes, location, direction, movement and non-manual components).

The "Mis Primeros Signos" dictionary [11] is the first children's dictionary in LSE with around 500 terms. It is a bilingual tool (sign language and spoken language) designed for uptake by Spanish educational system bilingual schools, targeting children aged from 3 to 8 years. The signs are portrayed by children educated in the bilingual methodology, making it a unique piece of material from this viewpoint.

These latest developments were designed as support for LSE courses and, while they are effective in this context, they cannot be considered full-blown dictionaries because they only include terms pertaining to the course lessons. Another important point to be taken into account is that they are not as mature as lexicons developed for other languages, which is reflected by the fact that there is either no or only limited two-way search. Additionally, taking into account the diversity of sign languages within each country, a dictionary needs to include a compilation of the geographical varieties of each sign to be considered complete. Finally, and this applies worldwide, a vital point is that development efforts are very much dependent on teams of developers outside deaf communities, which means that ideally these communities should be supplied with the tools they need to add new or modify existing signs. These four points summarise the primary goals of DILSE [12,13] as a bilingual LSE dictionary for the Spanish deaf community, using multimedia to render sign languages' visual properties.

3 Study and Compilation of Signs

A fundamental task to assure that a dictionary's content is both accurate and realistically reflects a language is to study this language and compile the most significant words. The words are the actual signs in DILSE. One key factor to be taken into account in the case of LSE is that there is no flexible and dynamic governing body to decide which words should be included in LSE and which should not. This, together with the deaf community's need to be able to communicate in their mother tongue (LSE), often calls for the creation of new signs. The development of such new signs is linked to the appearance of new concepts in everyday life (Internet, Web, mobile telephony, scanner, magnetic resonance, etc.). Because there is no governing body, a number of signs may emerge to represent the same concept. Before quality audio-visual communication systems (broadband videoconferencing, UMTS, etc.) became widespread, the dissemination of these new signs was confined to the region in which they appeared. This has led to what are termed geographical varieties of a sign.

A start was made to the work on compiling terms by thoroughly revising existing LSE dictionaries [8,9]. These dictionaries had several shortcomings. One weakness was that the format is not based on LSE as a self-contained language; it referred back to Spanish in all cases. Additionally, a sizeable number of the signs have fallen into disuse. Also examining some sign languages dictionaries put together in other European countries [14], we found that the signs were ordered by and search criteria were based on the hand shapes of each sign. This point was adopted as a criterion for building DILSE.

The last point was to compile the geographical varieties of LSE for each word. We analysed major Spanish dictionaries and second language teaching/learning material to make a preliminary selection of vocabulary. Then we contacted associations of the deaf from different Spanish regions and asked them to compile the signs that they were then using to translate the terms on the list we had put together. Signs that were commonly used in most regions were classed as "national".

4 Description of the Signs

In written languages, words are described by the letters of which they are composed. Obviously, we needed some way of describing a sign to be able to organise a bilingual LSE-Spanish dictionary. The initial hand shape before starting the sign and the movements made by the hands were chosen as a basis for classification, leading the CNSE [15] to establish the following characteristic factors as describing a sign:

- Active hand shape: This establishes the arrangement of the hand that does most towards forming the sign.
- Use of two hands: Some signs require the use of one hand and others involve both.
- Passive hand shape: The shape of the second, or passive hand, is also involved in the description of a two-handed sign.
- Body contact: This indicates whether or not the hands make contact with the body.
- Hand movement: The relationship between the shape of and the movement made by each hand plays a role in two-handed signs.

The different LSE words were then analysed and catalogued according to these characteristic factors, and the organisation of the sign-based search criteria was defined.

We adopted the initial active hand shape as the first criterion for starting a search, because it is the key feature for representing a sign. For this purpose, we first had to select the initial active hand shapes. We started by examining the two existing Spanish classifications [16,17] and how initial active hand shapes were arranged in other countries [18], which, although based on different languages, was a potential source of underlying criteria. At the end of this thoroughgoing process, we concluded that there were 49 active hand shapes for the signs entered in the dictionary, grouped, according to linguistic criteria, into 11 hand-shape families.

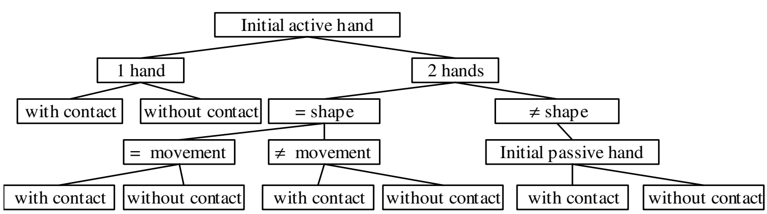


Fig. 1. Signs description

We were then able to define the order in which other sign classification criteria, such as whether or not there is body contact, the use of two hands, whether or not the movement and the hand shape of both hands is the same, were to be considered. These criteria can identify the properties of each sign, without the need of any previous linguistic knowledge (see Fig. 1).

5 Software Development

One of the essential aspects of system software development was to provide an interface tailored to the characteristics of users and the highly visual content of the information that was to be entered in the dictionary. Another vital point was to build a system that could include new functionalities and easily extend the dictionary contents, thereby giving the deaf community as much independence as possible and get them to identify with the tool.

5.1 Framework

It was decided to use the Constructor of Relations with the User (COREUS: *Constructor de Relaciones con el Usuario*) framework [19] to undertake the development project. This decision was made because this is a good framework for developing applications in which the interface is likely to undergo modifications at development time, because it keeps the application software separate from the interface. Also the user interface can be tailored to the specific features of the user. To this end, COREUS works on the principle of concept-based interfacing [20], according to which each item appearing in the interface represents a concept. The interface representation (visual, audio, haptic, etc.) will depend on parameters like the users' mother tongue, the application domain knowledge (in this case LSE), etc.

Because LSE is a visual language, it is only logical that part of the information required to consult the dictionary should be graphical. This led to the definition of an interface based primarily on icons and videos to set out the information required for LSE-based searching. Interface development is influenced by the information required to run such a search.

5.2 Search Algorithm

LSE-based searching depends above all on how the hands are positioned and the hand movements made to form the sign. Fig. 1 shows what sign classification criteria have been used to develop the search algorithm, and the interface that was to be used to gather information from users.

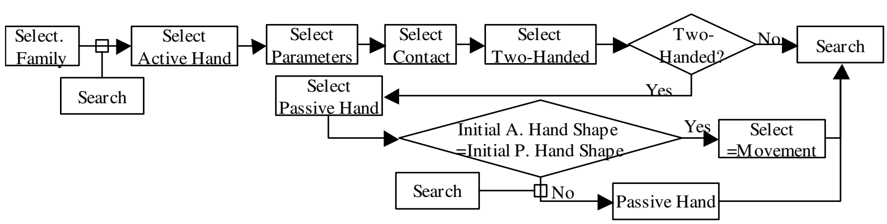


Fig. 2. Search algorithm

Fig. 2 is a diagram of the algorithm that is responsible for running the search process. In the first place, the signs from the database are filtered by the selected family. Then the user enters the appropriate sign parameters. Finally, a list of all the signs that match the selected search pattern is retrieved from the database.

6 Spanish Sign Language Dictionary (DILSE)

This section briefly discusses the two search types that can be run: search a word to get its respective sign and locate a sign to get its Spanish translation.

6.1 Searching by Spanish

Users type in a Spanish word to search for its respective sign or, alternatively, they can select the term from a word list. A results window is then displayed (Fig 4.b), showing all the information related to the located sign. This window displays the Spanish word, Spanish meaning and its context, Definition including an explanation in Spanish and a LSE video or an illustrative figure, Example of use, Lexical equivalences and Geographical varieties with links to other varieties. This window also includes an alphabetical listing of Spanish words, which can be used to search for other signs.

6.2 Searching by Sign

To run a search based on a sign, users have to go through a number of windows to gradually define the exact target sign. First, a window showing the initial active hand shape families is displayed (Fig 3.a). Users have to choose the family to which the sign hand shape belongs. Then another window requesting information on the sign properties, according to the above-mentioned criteria, is displayed (Fig. 3.b).

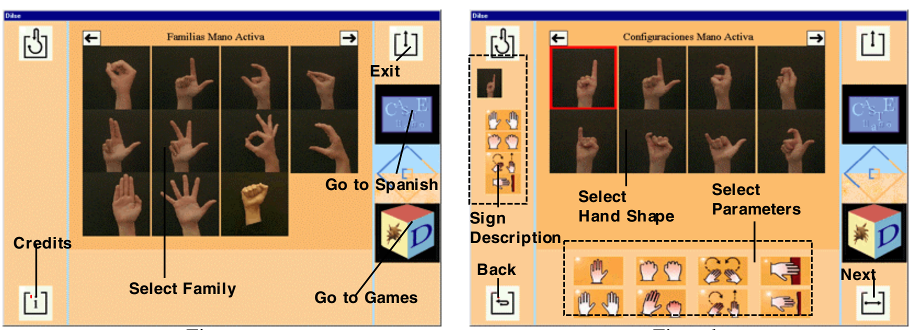


Figure a Figure b

Fig. 3. Selection of family and search parameters

The next window offers a selection of signs that share the selected properties (Fig 4.a). Upon selecting one of these signs, a list of the Spanish words that can be used to translate the sign, together with the geographical variety to which the sign belongs is displayed. Additionally, users can view a video showing the formation of the selected sign. Information on the sign parameters (initial active and passive hand shapes, sign location, palm direction, movement, hand direction and facial expression) can be gathered at any time.

If users select one of the words with its geographical variety, one last window (similar to the one viewed in Spanish-based searches) is displayed (Fig 4.b).

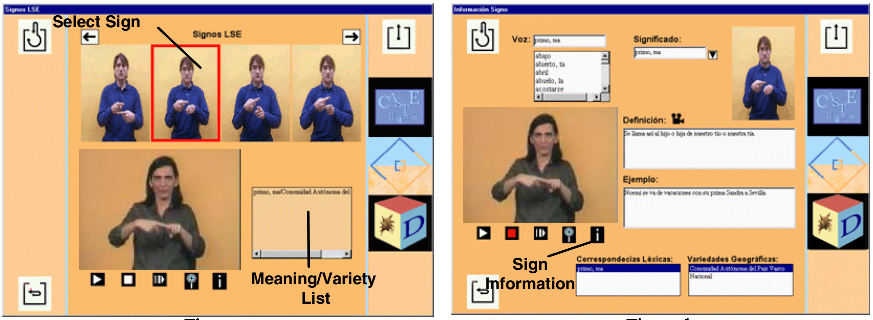


Figure a Figure b

Fig. 4. Sign selection and search results

6.3 Morphological Information

An additional DILSE utility is that it provides morphological information (related to classifiers and directional verbs). This information accompanies signs that undergo some sort of inflection that is dependent on the context in which they are used and has an effect on their formal structure. This means that a sign can have a different parametric structure depending on whether it is uninflected (for example, "avisar" -to warn-) or inflected ("os aviso" -I warn you-). This type of information is used for vocabulary learning through speech use contexts. An icon (representing a "C" if it is a classifier and a "V" if it is a directional verb) displayed in the results window indicates that morphological information associated with a sign is available. The classifiers window will contain a photo of the initial hand shape of the classifier and two videos (with the contextualised classifier and with an animation related to this classifier). The directional verbs window contains a list with some Spanish inflections of the selected verb and a video window showing the selected inflection in LSE.

7 Conclusions

DILSE is one of the first LSE-Spanish bilingual dictionaries for deaf people whose content can be accessed from the description of the LSE sign. It uses for this purpose an interface enabling users to select the graphical properties of the sign. The interface was developed on the basis of COREUS, which was used to build a platform for creating and maintaining sign language-based dictionaries. The DILSE dictionary of neologisms [13] and the DILSE basic dictionary [12] have been developed thanks to this platform.

Apart from the technological headway made thanks to platform development, another point worth considering is just how important it is to put together up-to-date and updateable electronic sign language dictionaries that are tailored to the characteristics and needs of this language and the user community. Exhaustive scientific research is a prerequisite for the standardization and full acceptance of sign language. Additionally, the only way of removing the communication barriers that the deaf community comes up against is to encourage the study, research and recognition of sign language. Bilingual dictionaries play an indispensable role in successfully achieving these goals.

The platform developed on the basis of the COREUS framework has opened the door to updating and developing sign language-based bilingual dictionaries leading to the complete integration of deaf people.

References

1. Oviedo, A.: Diccionarios de Lenguas de Señas. Grupo de Estudios de Lingüística de las Lenguas de Señas, Universidad de los Andes, Mérida, Venezuela. Available at: http://www.insor.gov.co/orientaciones_generales/diccionario.htm (last visited 7-3-2006)
2. Johnston, T.: Signs of Australia on CD-ROM. AUSLAN Dictionary. Deafness Resources, Australia (1997)
3. Johnston, T.: "AUSLAN Dictionary. A Dictionary of the Sign Language of the Australian Deaf Community. Deafness Resources, Australia (1989)
4. National Deaf Children's Society. Available at: http://www.ndcs.org.uk/ (last visited 7-3-2006)
5. Solina, F., Krapez, S., Jaklic, A., Komac, V.: Multimedia Dictionary and Synthesis of Sign Language. In: Rahman, S. M. (ed.): Design and Management of Multimedia Information Systems. Idea Group Publishing (2001) 268-281
6. Sign Language Online Dictionary and Lessons, ASL. Available at: http://handspeak.com, (last visited 7-3-2006)
7. British Language Online Dictionary, BSL. Available at: http://www.britishsignlanguage.com (last visited 7-3-2006)
8. Frigola, J.: Lenguaje de signos manuales. Ed. Dossat, 2nd ed., Barcelona (1998)
9. Pinedo, F. J.: Nuevo Diccionario Gestual Español. CNSE, Madrid (1988)
10. Rodrigo, J.: Signos 97-98. Diccionario de Lengua de Signos Española. Alicante (1998)
11. CNSE: Mis primeros Signos. Primer Diccionario infantil en lengua de signos. Spain (2005)
12. Martínez, F.: Diccionario de la Lengua de Signos Española (Diccionario Básico). CNSE, Spain (2000)
13. Martínez, F.: Diccionario de la Lengua de Signos Española (Diccionario de Neologismos). CNSE, Spain (2000)
14. Brien, D.: Dictionary of British Sign Language/English. Faber and Faber, London (1992)
15. Frutos, S., González, Á. L., Minchero, A., Nieto, M. J.: Diccionario de la Lengua de Signos Española (DILSE). Congreso Iberoamericano Iberdiscap 2000, Madrid, Spain (2000)
16. Rodríguez, M. A.: Lenguaje de signos. CNSE y Fundación ONCE, Madrid, Spain (1992)
17. Muñoz, I.: Aportación al estudio de la articulación de la Lengua de Signos Española en la Comunidad Valenciana. Undergraduate dissertation, University of Alicante, Spain (1996)
18. Prillwitz, S.: Hamburg Notation System (HamNoSys). Signum Press, Hamburg (1989)
19. Frutos, S., González, Á. L., Martínez, L. A., Montes, C.: Adapting Computer-Human Interaction in Real Time. In: Proc. of the IEEE International Conference on Systems, Man, and Cybernetics (SMC'98), San Diego, California, USA (1998) 1038-1043
20. González, Á. L.: Modelo para la generación y gestión en tiempo de ejecución de procesos de interacción hombre-máquina a partir de un lenguaje de especificación de relaciones con el usuario. PhD Dissertation, Universidad Politécnica de Madrid, Madrid, Spain (2003)

Improvements and Evaluations in Sign Animation Used as Instructions for Stomach X-Ray Examination

Kazunari Morimoto [1], Takao Kurokawa [1], and Syouhei Kawamura[2]

[1] Graduate School of Engineering and Science, Kyoto Institute of Technology, Matsugasaki, Sakyo-ku, Kyoto 606-8585, Japan
{morix, kurotak}@kit.ac.jp
http://media.kit.ac.jp

[2] Graduate School of Engineering and Science, Kyoto Institute of Technology, Matsugasaki, Sakyo-ku, Kyoto 606-8585, Japan
syouhei@ken.twinstar.jp

Abstract. To make sign language animation used as instructions for stomach X-ray examination easier for hearing impaired patients to understand, the authors identified problems in currently used sign animation and then created a new animation to solve these problems. The objective of this study was to compare how easy to understand the new animation was over previous animation by conducting evaluation tests. The evaluation tests were conducted with 32 hearing impaired persons. Results indicated that the new animation improved comprehension approximately 8% over previous animation.

Keywords: sign language animation, sign language interpretation, hearing impaired patient.

1 Introduction

There are far too few sign language interpreters for the number of hearing impaired persons. Because of the lack of qualified personnel, hearing impaired patients were concentrated in hospitals that retained an interpreter on their staff, but it was reported that this proved both physically and mentally stressful for the interpreters and even affected their health to the extent of shoulder-arm-neck syndrome, etc[1,2]. To solve this and other problems related to hearing-impaired persons and persons of normal hearing capability as well as between hearing impaired persons themselves, the authors are researching and developing a bi-directional Japanese-Sign language interpretation system.

This study focused on the communication between the X-ray technician and the patients undergoing a stomach X-Ray examination. In order to create instructions using easier to understand signal language animation, problems in existing sign language animation were improved and new animation-based instructions were created. Evaluation tests were then conducted to determine how easy to understand the new animation was. Comparisons were done with two existing sign language animations, the animation developed in this study and the signing of four actual interpreters.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 607 – 614, 2006.

2 Development of Sign Animation for Stomach X-Ray Inspection

2.1 Sign Animation 1

This sign language animation was created as the first step towards building an animated system for instructing hearing impaired patients during X-ray examination. In order to evaluate how easy the animation is to understand, evaluation tests were conducted with 8 persons total between 7 hearing impaired persons and 1 person with normal hearing capability. The 27 instructions given in Table 1 were used in the evaluation tests. The average correct response rate for all 27 instructions was low at approximately 30% (Uchikawa, 2000). After analyzing those results, some problems were pointed out in the animation (Uchikawa, 2001).

Table 1. Instructions for a stomach X-ray examination

No.	Instruction	No.	Instruction
1	Drink this medicine to inflate your stomach.	15	Relax.
2	Do not burp.	16	Extend your legs fully.
3	Drink all of the medicine.	17	I am going to move the table.
4	Hold the poles on your left and right firmly.	18	I am going to press this bar against your stomach.
5	I am going to lower the table.	19	Bring your right shoulder forward slightly.
6	I am going to lower your hands.	20	Hold the cup in your left hand.
7	Relax your muscles.	21	Hold the cup to your mouth.
8	Raise your right hip.	22	When I give the signal, drink from the cup.
9	Slowly set your hip back down.	23	Now drink.
10	Roll over on your stomach from the right.	24	Lift your chin slightly.
11	Face left.	25	Wiggle your chin.
12	Slowly roll over on your back.	26	Face right and lean forward.
13	Breath in.	27	That completes the examination.
14	Inflate your stomach and hold your breath.		

2.2 Sign Animation 2

Uchikawa et al. analyzed the problems in sign animation 1 (Uchikawa, 2001), and created a new sign animation 2 using a proposed method for making it easier to understand. Sign animation 1 was improved two separate times, the average correct response rate for all 27 instructions being about 30% after the first set of improvements and about 56% after the second set of improvements. The sign language of an interpreter had an average correct response rate of about 69%. The opinions on what was good, what was bad and what needed improving with regard to the sign language expression of this animation were categorized into three groups: those concerning sign language words, those concerning body gestures and those concerning the speed of sign language movements.

2.3 Improvements to Sign Language Animation

The following improvements were raised to solve the problems that were obtained from evaluation tests of sign language animation and make the animation easier to understand. (a) Inject rhythm into the sign language by adding pauses between words and working with the speed of hand movements. (b) Add new expressions using a more expressive animation model. (c) Add emotion. To satisfy these conditions, an attempt was made to change the animation so as to resemble as much as possible the sign language of an actual interpreter captured on video performing the 27 instructions for a stomach X-ray examination. In addition to improvements, detailed movements were improved so that the animation would more closely resemble that of an actual interpreter. For example, how to point one's finger when "point to the waist" and the adverb "slightly" were added. It was believed that resembling the sign language expressions of the animation to that of the interpreter would make the instructions for stomach X-ray examination easier to understand. (Uchikawa, 2001).

2.4 Improved Sign Language Animation

Instructions were created for X-ray examination by making some changes to existing sign language animation. Firstly, the same sign language expressionism as interpreters was applied to instructions for which the expressionism differed between the animation and the interpreter. Furthermore, attention was paid to the pauses between words, the emotion and speed of hand movements of the interpreters, and movements were made to best resemble those. Moreover, in order to express the sign language in the same way as interpreters, shoulder height and underarm aperture were added to the instructions in the animation. Figure 1 shows an example of instruction No. 5 "I am going to lower the table." The expression for "lowering" was made using the expression for "put your elbow by your inside," which was not possible in sign animation 2.

However, parts that did not resemble the sign language interpreter were a problem. They were caused by the animation program. These problems are summarized below.

(a) With words that take more time to say than form by hand movement, it was not possible to duplicate the expression of the interpreter. (i.e. "to inflate," etc.)

(b) How the fingers were attached to the model created an awkward doll-like impression (i.e., "stomach," "slightly," etc.)

(c) It was hard to change expressions (i.e., "relax," "hold," etc.) There were various reasons why the sign language of the animation did not resemble that of an interpreter. First of all, this animation does not calculate the number of frames required for finger movement until the word is expressed from the sign language code strings, therefore the program was designed to synchronize the mouth movement with the end of hand movement if the spoken word was short.

"I am going to lower" "I am going to lower"

Fig. 1. Example of improved sign language animation

3 Evaluation of Sign Animations by the Hearing-Impaired

3.1 Method

The evaluation tests compared the sign language instructions for stomach X-ray examination of sign animation 1, sign animation 2, the newly improved animation (hereinafter, "sign animation 3") and sign language of interpreters.

The subjects were sat before an LCD monitor and given a written explanation of the test method and the X-ray examination. It was explained to the subjects that the images displayed on LCD monitor were sign language instructions for a stomach X-ray examination. In addition, to make it easier for the subjects to picture in their minds an X-ray examination room, they were shown photographs of an actual examination room with an examination going on.

1) Evaluation of Comprehensibility of Individual Instructions

   The subjects read the sign language instructions shown on the LCD and wrote what they saw on an answer sheet that they were given in advance. This evaluated whether the response to each question was correct or not. The sign language instruction was shown twice and the subject responded separately for each.

2) Subjective Evaluation of Individual Instructions

   This part of the test evaluated how easy each instruction was understood. The following three responses were made available to the subjects. Each instruction was scored on a scale of 1 ~ 5. Evaluation items and scales were as follows.

   - Did you understand the meaning of the instruction?
     - 1. I did not understand the meaning. ~ 5 I understood the meaning.
     - 1. It was hard to understand. ~ 5 It was easy to understand.
     - Did the hands move at a suitable speed?
     - 1. Hand movement was slow. ~ 3. Hand movement was just right. ~ 5. Hand movement was fast.

Moreover, subjects were asked to provide their opinions on anything they noticed with regard to expression or what could be changed and where so as to make the sign language easier to understand.

3.2 Subjects

There were 32 subjects aged 20 and 79, all of which could read and write. There were 14 men and 18 women, amongst which their experience with sign language ranged from about 10 to 60 years or more. The 32 subjects were randomly divided into 16 pairs. Each subject responded to and evaluated the instructions in order from No. 1 to No. 27. The same instruction from the three animations and an interpreter exists. Among the four types of sing language expressions, the authors selected reasonable ones to make a group. Before testing began, the subjects were given a written explanation of the test and questionnaires were passed out. Answer sheets were then passed out so that subjects could write down what they interpreted from the sign language shown on the LCD. After the subjects completed filling in the information, descriptions of the sign language and subjective evaluation forms were carried out.

4 Results

4.1 Correct Response Rate to Instructions

Of the responses recorded by the subjects on the answer sheet, those that conveyed the meaning of the sign language was considered correct even if expressed differently from the instructions. In short, those judged to have the content of an instruction of an actual examination room was considered correct. For example, for instruction No. 3 "Drink all of the medicine," if the subject responded "Drink everything in the cup," the response was accepted because in an actual examination, the subject would be holding the cup when he/she is shown the instruction, so even if he/she interprets it as "Drink everything in the cup," it is likely that the subject will perform the action of instruction No. 3, which is to "Drink all of the medicine."

Results showed that the average correct response rate was about 75% for the interpreter's sign language, about 60% for sign animation 3, about 52% for the animation 2 and 30% for the animation 1.

The average correct response rate per instruction for the animation 3 was higher than that of sign animation 2 in 15 cases (Nos. 1, 3, 4, 5, 7, 8, 9, 11, 12, 13, 16, 17, 18, 19 and 20) with a difference in correct response rate of 2.5 ~ 50%. However, there was no difference in correct response rate between the animation 2 and sign animation 3 for 5 instructions (Nos. 6, 10, 15, 25 and 27). There were also 7 instructions (Nos. 2, 14, 21, 22, 23, 24 and 26) for which the animation 3 scored a lower correct response rate than sign animation 2 and exhibited a difference in correct response rate of 12.5 ~ 37.5%. The largest of those was instruction No. 24 at 37.5%, the next being No. 14 at 25%. The difference for the remaining 4 instructions (Nos. 21, 22, 23 and 26) was 12.5%.

In making sign animation 3, instructions (Nos. 3, 14 and 27) did not undergo major improvements. This was because a high correct response rate of 80% or higher was obtained in rating tests of animation 2. Of the 24 instructions excluding these 3 instructions, sign animation 3 scored higher than sign animation in 14 cases, the same in 4 cases and less than with the remaining 6 cases.

4.2 Relationship of Correct Response Rate and Time Structure

Figure 2 shows the relationship between the correct response rate of sign animation 3 and the total pause time per instruction. However, this does not include instructions Nos. 3, 14 and 27. The average pause time per instruction of all 24 instructions was 2.6 sec. The longest total pause time came with instruction No. 4 at 4.5 sec, while the shortest was with instruction No. 23 at 0.7 sec. As can be understood from Fig. 2, the correct response rate is higher than 40% with instructions which total pause time per instruction was longer. The curve shows upward curve (higher response rate with longer pause time) is the evident tendency. However, there were instructions of a short total pause time that had a low correct response rate as well as others with a high correct response rate. Therefore, a low correct response rate cannot be attributed to a short total pause time.

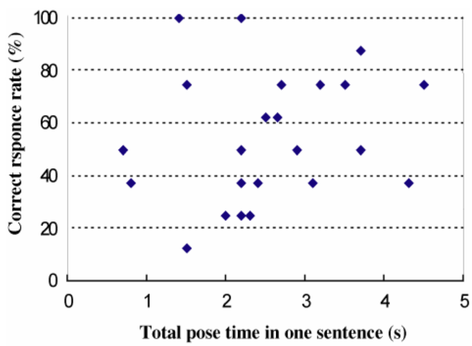


Fig. 2. Relationship of correct response rate to total pause time per instruction in sign animation 3

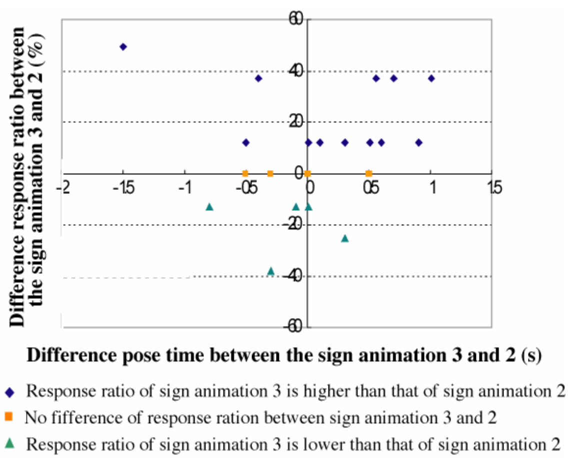


Fig. 3. Relationship of difference in correct response rate between sign animation 3 and sign animation 2, and difference in pause timer per instruction

Figure 3 shows the relationship between the difference in correct response rate between the animation 3 and sign animation 2, and the difference in total pause per instruction between the two. Looking at sign animation 3, which had a higher correct response rate than sign animation 2, 8 instructions (Nos. 1, 6, 7, 8, 10, 11, 15, 17, 18 and 19) had a long total pause time per instruction, while the total pause time for 1 instruction (No. 4) did not change. Also, the pause time was shorter in sign animation 3 than in sign animation 2 in 3 cases (Nos. 5, 6 and 17). Yet, though the pause time per instruction was shorter (-1.5 sec) than in sign animation 2, sign animation 3 had 17 instructions that improved the correct response rate by 50%. With the exception of this instruction, a trend was seen in that the longer the pause time, the higher the correct response rate, whereas the shorter the pause time, the lower the correct response rate.

4.3 Subjective Evaluations

Looking at the average results per instruction, the meaning was more easily understood and it was easier to divide words in the order of the interpreter, sign animation 3, sign animation 2 and sign animation 1. As for the suitability of hand speed, the order from best to worst was the interpreters, sign animation 3, sign animation 2 and sign animation 1. The correct response rate was lower than that for sign animation 2 with instructions Nos. 4, 8, 21, 23, 24 and 26. With the exception of No. 23, it was harder to understand the meaning and divide up words than in sign animation 2 with instructions Nos. 4, 8, 21, 24 and 26. Moreover, the hand movements of instruction No. 26 were faster with sign animation 3 than with the other three signs.

5 Discussions

After improvements to the sign language animation, the correct response rate and subjective evaluations got better in general, but with some words, the correct response rate still did not improve from previous levels (sign animation 2). More specifically, they were, 5 “I am going to lower the table,” 6 “I am going to lower your hands,” 8 “Raise your right hip,” 16 “Extend your legs fully,” 18 “I am going to press this bar against your stomach,” and 25 “Wiggle your chin.” With all of these instructions, the correct response rate was under 40%, so it is still far from being called an easy-to-understand animation. These instructions are conveyed with body gestures; if the body gestures are not understood, it is unlikely that the instruction will be understood.

The perceivable reasons why sign language animation cannot express the body gestures of an interpreter are as follows. The animation has trouble expressing subtle movements made by the interpreter. Moreover, the emotion and expression of an unfamiliar animated model are unnatural and thus harder to understand than a human interpreter. Some of the subjects commented that the hand movements of the animation were awkward and unnatural.

6 Conclusions

This study aimed to create sign language animation for instructions for a stomach X-ray examination by focusing on the application of a sign language-Japanese machine translation system. With earlier instructive animations, the highest evaluation was only 49%. This was because earlier Japanese to sign language interpretation system could not express emotion, mistakes were made in the position and mode of expression of the sign language, and movement was monotone. An animation was created with improvements to all of these areas and used in rating tests with hearing-impaired persons to see how easy to understand it was. Tests compared two earlier versions of animation, the improved animation and an interpreter. Results indicated an average correct response rate of about 56% for the new animation, which was higher than earlier animations.

Acknowledgements

This study was made possible by a Grant-in-Aid for Scientific Research (Basic Research B-1, No. 14370122 and B, No. 16300029) of the Japan Society for the Promotion of Science. The authors would like to express their appreciation for this support.

References

1. Kitahara, T., Taoda, K., Watanabe, S., Sato, S., Nishiyama, K.: *Are Deaf People Inhibited from Seeking Medical Services – Surveys on Deaf People and a Hospital Due to Lack of Sign Language Interpreters*, Bullet. Social Medicine, 14, 103-107 (1999) (in jpn)
2. Kurokawa, T., Morichi, T., Watanabe, S.: *Bidirectional Translation between Sign Language and Japanese for Communication with Deaf-Mute People*, Human Computer Interaction: Software and Hardware Interface, Elsevier, 1109-1114 (1993)
3. Nakata, Y., Morimoto, K., Kurokawa, K., Kitahara, T., Taoda, K. et al.: *Proposal on Stomach X-Ray System with Sign Language Animation for Deaf People*, Proceedings of Kansai-Branch Conference in Ergonomic Society of Japan, 109-112 (1998) (in jpn)
4. Morimoto, K., Uchikawa, K., Ogawa, K., Kurokawa, K. et al.: *Evaluation of Japanese Sign Language Animation for Hearing-Impaired People in Stomach X-Ray Inspection*, Human Interface, 2(5), 174-179 (2000) (in jpn)
5. Morimoto, K., Kurokawa, T., Isobe, N., Miyashita, J.: Design of Computer Animation of Japanese Sign Language for Hearing-Impaired People in Stomach X-Ray Inspection, Computers Helping People with Special Needs, Springer, 1114−1120 (2004)

Speech Technologies in a Computer-Aided Speech Therapy System

András Kocsor and Dénes Paczolay

University of Szeged Department of Informatics
6720 Szeged, Árpád tér 2
{kocsor, pdenes}@inf.u-szeged.hu
http://www.inf.u-szeged.hu

Abstract. The hearing impaired have always had difficulties learning to speak because their auditory feedback is either damaged or missing. The SpeechMaster software package provides real-time visual feedback as a substitute for this. Within the package the forms of the feedback are clear and simple. For instance in the first phase of vowel learning the software uses an effective phoneme recognizer providing real-time visual feedback. In this case flickering letters indicate correctness where the brightness of the letters is proportional to the output of speech recognizer. These unambiguous solutions help the hearing impaired to learn the correct association between the phoneme - grapheme pairs or the connection between their own articulation and the speech signal they produce. Also, with the aid of the computer, children can practice without the need for the continuous presence of the teacher. This is a significant step in the education of the hearing impaired as their traditional therapy includes a long and tedious fixation phase. Furthermore, the use of computer exercises, which are popular with children, speeds up the learning process.

1 Motivation

In the history of the education of the hearing impaired many aids and methods have appeared - some successful and some less so - which aimed to improve their pronunciation [3]. The teachers of the hearing impaired have always felt that the target group's pronunciation problems could be significantly (if not radically) reduced by a solution that would allow the hearing impaired to monitor and correct their own pronounced sounds. In the case of the very seriously impaired, near-normal hearing can be developed following a successful cochlea implant operation. After this, audible speech can be obtained following a period of intensive learning.

It is known from experience [5] that the training of the utterance of vowels is more difficult than that of consonants because their phonation is not so easy to explain. The key feature of the therapy of the hearing impaired is the refined pronunciation of vowels in order to attain articulate/intelligible speech. It would be a real help in therapy if the computer was able to provide an objective rating

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 615–622, 2006.

of the quality of the uttered vowels. If it is reliable and matches the subjective opinion of the therapist then it will relieve teachers of the burden of the tedious work they have with traditional therapy. This is a major problem because of the general lack of therapeutists. The main objective of the hearing impaired is not to speak with a perfect accent, but rather to speak intelligible as soon as possible. There is no perfect speech sample for the objective computer decision, but human hearing is tolerant as well. Therefore, by using large speech databases [7], the goal of the computer-aided speech therapy system is attainable; that is, with statistical machine algorithms, speech - which is intelligible for human hearing - is also recognizable.

2 Related Work

Up till now "Box of Tricks" was the only speech therapy software package available for the Hungarian language. This software was developed by RCS Bt. in cooperation with the Budapest University of Technology and Economics (Department of Telecommunications and Media-Informatics, Laboratory of Speech Acoustics). The development was carried out in a SPECO project for four languages (Hungarian, English, Swedish, Slovenian). "Box of Tricks" provides some playful and pleasant drills for learning the pronunciation of speech sounds. Fig. 1 shows the screen for learning the pronunciation of vowels and spirants. The drill displays the spectrum-like curve of the pronounced phonemes in real-time. The darker zone (track) is the acceptance zone - which is found from the average of curves calculated over a large speech database. The goal is to produce a pronunciation where the curve fits the above-mentioned zone. If we pronounce another phone then the humps of the curve shift, but their position is hard to control. Fig. 2 shows the drills of the pronunciation of the phones embedded in a word and in a sentence. The upper figure is the simplified spectrogram of the etalon utterance, while the lower one is the current utterance. The pronunciation is good when the two simplified spectrograms are similar, but what do we mean by "similar"? There are many more software package available for the English language

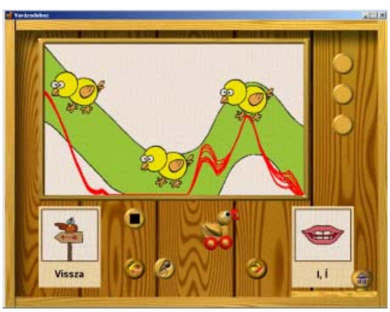


Fig. 1. Box of Tricks: Learning the pronunciation of isolated vowels

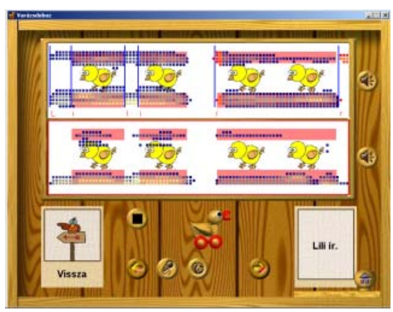


Fig. 2. Box of Tricks: Learning the pronunciation of vowels uttered in a sentence

Fig. 3. Main page of SpeechMaster software

Fig. 4. Vocalization drill

(e.g. VideoVoice, SpeechViewer, Dr. Speech), but they do not usually provide really clear real-time visual feedback, and are not suitable for the Hungarian language. More often than not these programs use advanced sound processing methods, but do not apply modern automatic speech recognition. Without automatic speech recognition the users need to make an additional comparison or evaluation and this is difficult for young learners. Lastly, we would like the speech therapy to be available without charge. The price of the "Box of Tricks" is 200 euros for home users and 800 euros for therapists. The prices of other software packages we checked typically cost thousands of euros.

3 The SpeechMaster Software Package

The SpeechMaster software package was developed for speech impediment therapy and teaching reading with the financial support of the Hungarian Ministry of Education. Fig. 3 shows the main page of the software. The program is freely available at the Internet site: `http://www.inf.u-szeged.hu/beszedmester`

In the therapy of the hearing impaired one of the key problems is how to deal with the lack of proper auditive feedback, which of course, impedes the development of intelligible speech. The idea is then to make the vocal sounds visible for the hearing impaired because so they are able to check their pronunciation by sight, just like others do with their hearing. The importance of our software package is that it puts into practice a speech evaluation that is based on automatic (machine) speech recognition, which effectively supports the hearing disadvantaged in the acquisition of articulate speech. The speech therapy of the hearing impaired traditionally requires enormous patience and the continuous presence of the teacher since, during the fixation of the correct sound-formation, a large amount of repetition and correction by the teacher are both needed. This so-called automation process is significantly speeded up and simplified using our software, and also allows the students to practice on their own and with the teacher.

The aim of the reading-teaching part of the software is to help children learn, more quickly and easily, the correct phone-grapheme and grapheme-phone

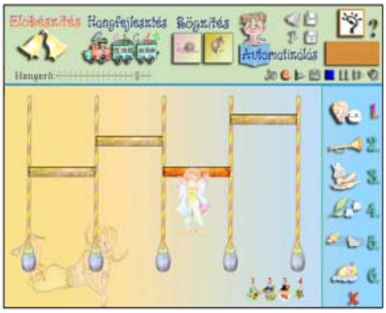


Fig. 5. Drill for loudness control

Fig. 6. Drill for rhythm

associations using the computer as a means of motivation, and by using automatic speech recognition methods. In addition to the general use of speech therapy, the program can also be used in the training therapy for the partially impaired and can also be a great help in dyslexia therapy and the treatment of certain speech impediments.

3.1 The Playful Sound Formation Exercises of SpeechMaster

In speech impediment therapy, at the beginning of the development of oral competence, it is recommended that young children concentrate mainly on their own voicing. This is supported by the creation of the following drills as part of SpeechMaster. In every drill, skill and acceptance levels as well as threshold values are adjustable with a potentiometer.

Fig. 4 shows the vocalization drill. The idea is to move the car as fast as possible with the help of clearly formed and pronounced, sustained, voiced sounds. The speed of the car is proportional to the 'periodicity', which is estimated via the Autocorrelation Function [6] (ACF) and smoothed by a moving average.

Fig. 5 shows the drill for loudness control. In this drill with the loudness of speech the degree of the clown's elevation can be influenced. The clown should be kept at the marked level for a length of time previously defined in the settings. The estimation of loudness is made based on the energy of the received signal, which is again smoothed by taking a moving average. The calculated energy and the clown's elevation are linearly proportional.

Fig. 6 shows the drill for rhythm. A jumping frog's footprints in the sand display the rhythmic scale of sounds uttered by the teacher. The student has to repeat the rhythm so the frog should jump on the footprints just made. During the drill the length of the frog's jumps can also be manipulated by the loudness of speech. The calculation of the loudness is the same as that described above, the applied scale also being linear.

Fig. 7 shows the drill for pitch control. The aim of the drill is to help children control the pitch of their voice. In the first part of the drill one plants flowers in the top and bottom rows according to the pitch of the voice. Afterwards the person practicing has to water the flowerbed by trying to repeat the previous series

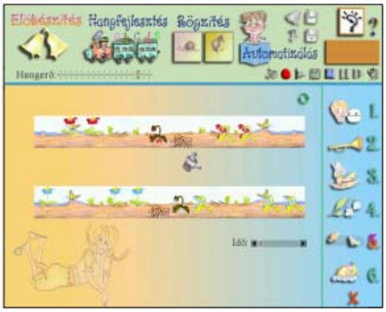


Fig. 7. Drill for pitch control

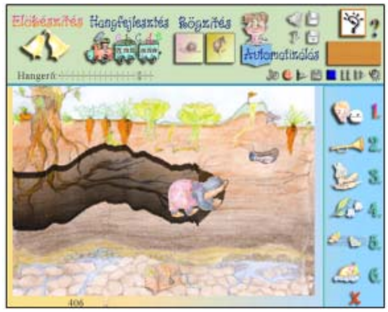


Fig. 8. Fine pitch control

of low and high sounds. The fundamental frequency estimation is performed by the ACF, and then the low/high pitch decision for the whole segment is made based on the majority voting rule.

Fig. 8 shows the drill for fine pitch control. The drill seeks to help the acquisition of continuous fine pitch control. First, by changing the pitch of a sustained sound, a mole digs a route in the soil. In the case of high pitch it digs closer to the surface and in the case of low pitch it digs deeper. During a practice session a smaller mole has to be led through this route. The exercise is divided into two major parts: defining the base voice and the coordination of the mole. Pitch definition is also done with the aid of the ACF and smoothed by a short distance moving average. The curve drawn by the mole and the fundamental frequency functions formed are logarithmically proportional.

3.2 Learning the Pronunciation of Vowels

In this phase the student learns the articulation of several vowels. In SpeechMaster the role of effective real time vowel recognition is essential. The software has many feedback configurations and allows the half speed replay of recorded speech sounds. The student can use a web-camera as a ”phonetic mirror” to validate his/her own articulation or compare one with the utterance of the teacher.

In Fig. 9 the main task of speech impediment therapy is shown. Real time visual feedback helps improve the articulation of the student. In Fig. 10, SpeechMaster shows the degree of phoneme recognition in a chart.

We separated our vowel database for men, women and children speakers, as well as isolated pronounced vowels and ones uttered in a word. The signals were processed in 10 ms frames, the log-energies of 24 critical-bands being extracted using FFT [2] and triangular weighting. In SpeechMaster we applied 2x3+1 three-layer Artificial Neural Networks [1] . The feed-forward MLP networks was trained with the backpropagation learning rule. Currently, the accuracy of the real-time vowel recogniser present in SpeechMaster is over 95%.

The output of the ANN is a real number between 0-1 for each vowel and approval level can be defined on this scale. The horizontal lines on the diagrams (Fig. 10) are the approval level, which are separately adjustable for each vowel.

Fig. 9. Learning the pronunciation of vowels

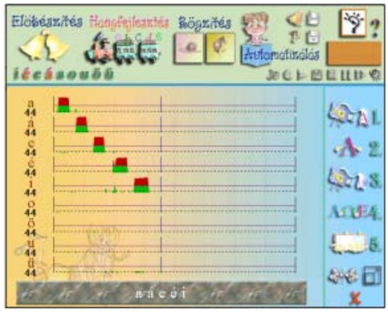


Fig. 10. Chart of phoneme recognition

The colour of the diagrams above the approval level changes from green to red. The therapist can save the approval levels for each students separately in their personal user profiles. These profiles ensure a real sense of achievement, motivation and continuous improvement. The teacher can record a standard utterance of the text which is typed in the lower text field. The children can replay the recorded standard utterance just by pressing a button. Then the therapist can save and replay the utterances of the children into their profile. These utterances are stored in a separate directory in chronological order. If the therapist saves a group of selected utterances week after week, then he can analyse the speech status of each child.

3.3 Fixation of the Pronunciation of Vowels in Words

In this phase the student practises the pronunciation of vowels in words. The software provides many word and image pairs grouped by grapheme. The vowel recogniser is identical with the one used in the previous phase, hence one can examine the utterances using other visual feedback (see Fig. 11).

Similar to theprevious drill, the teacher can record a standard utterance for each word-picture pair and save the utterances of the children. The teacher can decide to use the approval levels or not. If these levels are applied then the grapheme on the picture only appears if the intelligibility of the corresponding phoneme is above the level set.

3.4 Automation of the Pronunciation of Vowels in Sentences

In this task the impaired children practise the correct pronunciation of the vowels by reading a longer text (see Fig. 12). The text can be split in smaller units, which are defined by commas. The unit to be read is highlighted in pink, and the vowels in it are highlighted in red. When the student reads the words the cursor (vertical black line) also moves through the text. If the detected phoneme matches one to a good level then the letter's red colour disappears. The given part is completed when all observed phonemes on the screen are no longer marked.

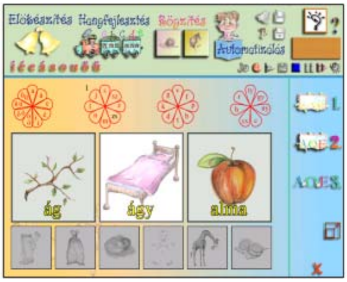


Fig. 11. Fixation the pronunciation of vowels in words

Fig. 12. Automation of the pronunciation of vowels in sentences

After the end of the selected text the software automatically moves on to the next part. The level of acceptance can be varied for each vowel separately. The therapist can freely type sentences into the text area, but the drill contains 10 fixed and 6 variable texts that are stored in the profile.

In this drill we combined the well-known dynamic time warping method with our stochastic segmental recognition technique, and also added features to the model that made it run faster and more smoothly. We arrived at the latter after a lot of experimentation.

4 Results Achieved with SpeechMaster

The design of the software package was completed in 2004. From the aspect of usability, our SpeechMaster system applies the same functions as the Eriksson's article recommended [4] in 2005.

The speech impediment therapy part of SpeechMaster was tested at the School for the Hearing Impaired in Kaposvár. For 5 months the software was used in the therapy of deaf, hard of hearing, implanted children and those who had

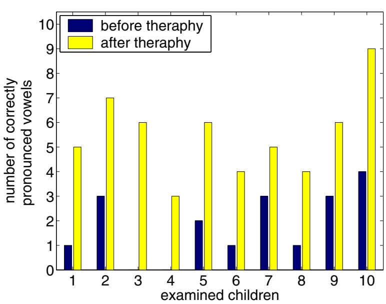


Fig. 13. In the case of 10 children, following therapy, the number of correctly pronounced vowels significantly improved

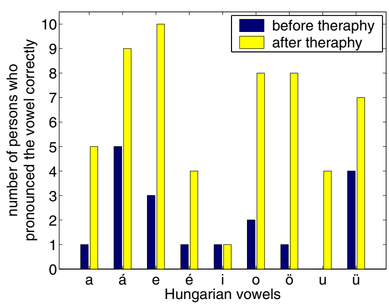


Fig. 14. The number of correctly pronounced vowels before and after therapy with the SpeechMaster software package

a speech impediment. Previously, all children attended 1-2 years of traditional therapy. After using SpeechMaster they all showed a significant improvement: the number of articulate vowels of the deaf children grew from 2-3 to 6-7. Following therapy each children pronounced vowels more clearly, and the number of various mistakes (e.g. nasalty) became noticeably fewer. The results of this are shown in Figs. 13 and 14.

5 Summary and Future Work

The speech-oriented parts of SpeechMaster in the fields of computer-aided teaching are innovative because, due to the speech interface, the interaction between the computer and user has become more human. The learning/therapy is the 'friendly' interaction of the child/hearing impaired.

During the development of the software package we found high quality solutions to several speech technology and signal processing problems. In the SpeechMaster software package we provide solutions to problems like sound and word recognition, as well as to the various types of speech evaluation. However, the opportunities of reading evaluation are only partially utilized and merely limited to speech impediment therapy. Hence we are planning further investigations in this area of research. Our intention is to develop a more complex and flexible automatic reading evaluation system. Moreover, it is quite conceivable that our software package could be adapted to other languages so non-Hungarian therapists could use it as well.

References

1. C. M. Bishop. *Neural Networks for Pattern Recognition.* Oxford Univ. Press, 1995.
2. E. Oran Brigham. *The Fast Fourier Transform.* Prentice-Hall, Inc., 1974.
3. Leen Cleuren. Speech technology in speech therapy?, 2003.
4. E. Eriksson, O. Bälter, O. Engwall, A.-M. Öster, and H. Kjellstörm. Design recommandation for a computer-based speech theraphy system based on end-user interviews. In *Tenth Int. Conference on Speech and Computers*, pages 483–486, 2005.
5. G. Hégely and A. Kocsor. A vizuális beszédértékelés alkalmazásának magyar vonatkozású történeti áttekintése a hallássérültek beszédoktatásában. *Alkalmazott Nyelvtudomány*, 2005.
6. L. R. Rabiner and R. W. Schafer. *Digital Processing of Speech Signals.* Prentice-Hall, Inc., 1978.
7. Gy. Sejtes, J. Bácsi, and A. Kocsor. Egy nem csak fonológiai adatbázis specifikációja. To be published, 2005.

Automatic Synthesis of Training Data for Sign Language Recognition Using HMM

Kana Kawahigashi[1], Yoshiaki Shirai[2], Jun Miura[1], and Nobutaka Shimada[1]

[1] Department of Mechanical Engineering, Osaka University
[2] Department of Human and Computer Intelligence, Ritsumeikan University

Abstract. The paper describes a method of synthesizing sign language samples for training HMM. First face and hands regions are detected, and then features of sign language are extracted. For generating HMM, training data are automatically synthesized from a limited number of actual samples. We focus on the common hand shape in different word. The database hand shapes is generated and the training data of each word is synthesized by replacing the same shape in the database. Experiments using real image sequences are shown.

1 Introduction

The paper deal with recognition of sign language from a video sequence. Features of sign language consists of the position, velocity, and shape of hand regions. Sign language words are trained and recognized using Hidden Markov Model (HMM) [1]. For building reliable HMM, many training samples are necessary. However, collecting many samples from different subjects is difficult. Therefore we propose a method of synthesizing training data automatically from a limited number of actual samples.

In speech recognition, speech samples are synthesized from phoneme units [2]. In CG animations of sign language, sign sentences are generated by combining the patterns of words [3]. However, there is no methods for synthesizing training data of sign language. Here, we focus on the hand shape. The database of sign language word and its constituent contour shapes is generated from sign language images. The training data of each word is synthesized by replacing the contour shapes with the same shape in other words.

2 Feature Extraction

2.1 Extraction of Hand Regions

Face and hand regions are first extracted using a model of skin color, the range of which is determined from the initial image.

When hands and face regions overlap, they are separated using the previous an the succeeding frames. First, the image of the face and hands just before overlapping are saved as templates. Assuming that those images does not change, hands and face regions are extracted by template matching.

Next, assuming that hand shape changes during overlapping, the hand region is extracted. The image of face and hands just after overlapping are saved as templates and the regions are similarly extracted by template matching.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 623–626, 2006.

Then the timing of the shape change is determined comparing the degree of matching in the forward backward template matching.

2.2 Extraction of Hand Features

From the extracted face and hands region, the positions, the velocities and the shapes of hands are obtained.

Because a small position difference of a hand near the face or a small movement near the face is often important, the positions and velocities of hands are:

- The log-distance between the face and each hand region.
- The change of the log-distance between the face and each hand region.
- The direction of each hand from the face region.
- The change of the direction of each hand from the face region.

Because, for two hand gestures, the relative position is important, the relative position of the right hand to the left hand is also included in the features

The shape features of the hand include:

- The number of protrusions of the hand region.
- $\{u(1-r), v(1-r)\}$, where
  (u, v): The x and the y component of the principal axis of the hand region.
  $r(0 < r < 1)$: The degree of circularity of the hand region.

Because the direction of the principal axis of a circular region is unstable, weight $1 - r$ is multiplied to decrease the value of the direction.

3 Generation of Initial HMM

Initial HMM is generated from states corresponding to the motion of hands. The image sequence is first segmented into static and moving periods. Even if a hand moves slowly, the corresponding static period is segmented into two if the moving direction changes significantly. Moreover, if the distance between the face and a hand is small and the direction of the hand from the face changes significantly in a static period, the period is also segmented into two.

The means and the variances of features in each state are calculated to create an initial model.

4 Synthesis of Training Data

4.1 Making a Hand Shape Database

Fukuda et al. proposed a classification of the hand shape used in sign language [4]. This classification is based on the finger alphabet of the Japanese syllabary. In order to deal with words which are not expressed by those alphabet, we add three hand shapes.

Moreover, since the view of a hand in images changes depending on a hand position and pose, we further extend the classification to include typical variations of the same hand shape. Figure 1 shows the extended classification.

Because the shape is important while a hand is not moving, the static hand shape features are saved in the database.

Fig. 1. Classification of hand shape

4.2 Synthesizing New Training Data

In order to synthesize a training data of a word, the degree of circularity is checked to find static periods. Then, for each hand shape, the database is searched for the similar hand shape. If it is found, the shape features replace the original features, while the other features such as the position and the motion are kept unchanged.

5 Experimental Results

Experiments of recognition is performed using the HMM generated from original image sequences and the synthesized data. In the experiments, images of two persons are used. For each word, three sequences of images are obtained and the total of six samples are obtained.

Excluding words whose images are not successfully processed, 14 words expressed by both hands and 21 words expressed by one hand are used. First, only actual data are used for training, where five of the sequences are used for training and the rest is used for the test. By changing the sequence for the test, six experiments are performed.

Next synthesized data sequences are added for training and similar six experiments are performed. Experimental results are shown in Table 1. Recognition rates of the experiments using only actual data and the experiments using synthesized data have little differences.

Before adding synthesized data, the word “like” of subject A is misrecognized as the word “red”, and the word “red” of subject B is misrecognized as the word “dislike”. By adding synthesized data, they are recognized successfully.

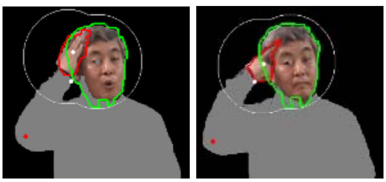
Table 1. Experimental results

(a) Subject A

	both hands		one hand	
synthesized data addition	before	after	before	after
the number of success	52/52	52/52	57/63	57/63
recognition rate	100%	100%	90.5%	90.5%

(b) Subject B

	both hands		one hand	
synthesized data addition	before	after	before	after
the number of success	50/52	50/52	57/63	58/63
recognition rate	96.2%	96.2%	90.5%	92.1%



(a) black (b) head

Fig. 2. Words of similar hand features

On the other hand, the word "head" shown in Figure 2(a) of subject A is misrecognized as the word "black" shown in Figure 2(b). Because the hand shape view of these words are similar, it is easily misrecognized. Even by adding synthesized data, the difference is ambiguous. We need better features to discriminate them.

One of the reason of misrecognition in either case is that the test data is far from training data. It is necessary to collect hand shape data for more words in order to synthesize a variety of training data.

6 Conclusion

We proposed a method of automatically synthesizing of training data using a hand shape database for sign language recognition. Hand shapes in sign language words are classified and training data are synthesized by replacing hand shape features.

Although experiments with a small number of sign language words, the effect of using synthesized data is not prominent, the method seems promising if it is applied to larger data set.

The future problem is in addition to extension to a larger data set, application of the method for various persons.

References

1. J. Kinscher, H. Trebbe, "The Munster Taging Project - Mathematical Background", *Arbeitabereich Linguistik, University of Munster*, D-58149 Munster, 1995.
2. T. Toda, H. Kawai et al, "Unit Selection Algorithm For Japanese Speech Synthesis Based on Both Phoneme Unit and Diphone Unit", *Proc. Int. Conf. Acoustics, Speech and Signal Processing (ICASSP2002)*, pp. 3088-3091, 2002.
3. H. Sagawa, M. Ohki et al, "Pattern Recognition and Synthesis for a Sign Language Translation System", *Jounal of Visual Languages and Computing* Vol. 7, No. 1, pp. 109-127, 1996.
4. Y. Fukuda, H. Kimura et al. "Expression through hand and finger actions used in 'words in continuous signing': A study in reference to the dialogue in Japanese Sign Language by the deaf persons", *Technical Report of IEICE*, Vol. 97, No. 586, pp. 81–85, 1998 (in Japanese).

Communication Supporting System in a Classroom Environment for the Hearing Impaired

Yoshinori Takeuchi[1,2], Yudai Sakashita[3], Daisuke Wakatsuki[4], Hiroki Minagawa[4], and Noboru Ohnishi[3]

[1] Information Security Promotion Agency, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, 464-8603 Japan
takeuchi@nagoya-u.jp
[2] Bio-Mimetic Control Research Center, RIKEN, 2271-130, Anagahora, Shimoshidami, Moriyama-ku, Nagoya, 463-0003 Japan
[3] Graduate School of Information Science, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, 464-8603 Japan
{sakashita, ohnishi}@ohnishi.m.is.nagoya-u.ac.jp
[4] Faculty of Industrial Technology, Tsukuba University of Technology, 4-3-15, Amakubo, Tsukuba, Ibaraki, 305-8520 Japan
{waka, minagawa}@a.tsukuba-tech.ac.jp

Abstract. In this paper, we propose a communication support system in a classroom environment for the hearing impaired. This system detects the questioner, monitors his/her sign language with a video camera, and displays the video on a monitor at the front of the lecture room. Although other students cannot see the questioner's sign language directly, they can see the questioner on the monitor. The system locates the questioner by detecting a raised hand. Once the system locates the questioner, it zooms in by controlling the direction and the zoom parameters of the camera and captures his/her image. We implemented the system and conducted an experiment during a real lecture. As a result, we achieved a raised hand extraction rate of 70%.

1 Introduction

Most hearing impaired students communicate with sign language when they ask a lecturer a question in class. Sign language is a visible language, and, as a result, some other students may not be able to see the questioner. Figure 1 illustrates this situation. It is likely that the lecturer may be asked the same question many times, since students may not realize the question has already been asked.

To solve this problem, we developed a system that monitors the speaker and displays him/her on a monitor placed at the front of the room, as shown in Fig. 2. Other students can see the monitor, which allows them to engage when a question is posed. A student who wants to ask a question shows his/her intention by raising a hand. The system locates the questioner by detecting his/her hand. Once the system detects the hand, it zooms into the questioner. Then, the system

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 627–634, 2006.

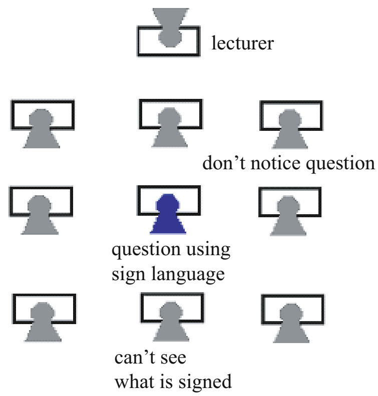


Fig. 1. Lecture for hearing impaired people

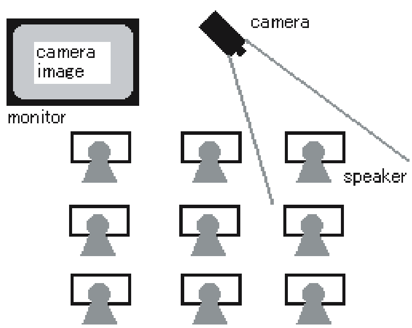


Fig. 2. Automatic monitoring system

monitors the questioner's sign language with a video camera and displays footage on the front monitor. In this paper, we propose a method for detecting the questioner and controlling the zoom camera. We also implemented the system by a pan-tilt-zoom camera and conducted an experiment in a real lecture for hearing impaired. Finally, we discuss the evaluation of the experimental results from the real lecture.

2 Detecting a Raising Hand

In this lecture, a student who wants to ask a question will show his intention by raising his/her hand. In this section, we propose a method for detecting a raising hand. Our system must run in realtime so we must avoid time-consuming computational tasks.

Our detection method consists of two parts. First, we find the location of each student's head. We find students' head locations using background subtraction, skin color extraction, face area extraction, and hair area search. Next, we detect a raised hand around the head location. We describe the details in the following section.

2.1 Estimating Head Location

In order to detect the questioner from the rest of the students, the system must know the location of each student. A student raises his/her hand around his/her head. Therefore, the system needs to know the head location of each student. This is useful for reducing the area required for hand detection and for monitoring student's sign language actions. We describe the estimation procedure steps in the following section.

Background Subtraction. First, we pinpoint the students' region using background subtraction [1]. We take a background image in advance, and subtract it from the current frame. Figure 3 illustrates an example of this method.

Fig. 3. Example of background subtraction

If we applied this method to a real lecture, it would fail due to changes in illumination caused by turning the lights on/off or opening/closing curtains. Therefore, we must update the background image during the lecture. The background image is updated using Eq.(1).

$$B_{t+1} = (1 - \alpha)M_t + \alpha B_t, \quad (1)$$

where B_t is the background image at time t, M_t is the background region extracted from the current image, and α is a constant $(0 \leq \alpha \leq 1)$.

Skin Color Extraction. In order to locate the face region, we extract the skin color region from the current frame [2,3]. We use the YUV color space; Y represents intensity, while U and V represent chromaticity. We extract the skin color region using the U and V values.

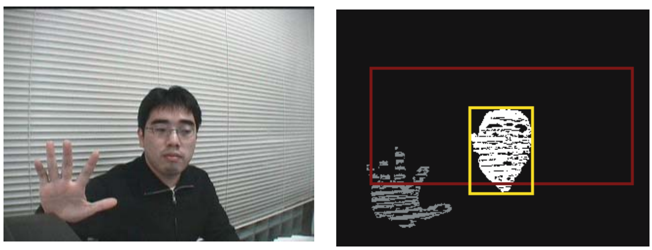
In this paper, we use a simple extraction method. We make a UV region which represents skin color in advance. Skin color is then extracted if the UV value is inside the skin color region. The UV value of skin color is slightly affected by illumination so we make skin color regions under several different illumination conditions.

Facial Region Extraction. We identify the facial region candidates from the skin color region [4]. We eliminate regions whose areas are too small or too big, or whose aspect ratios (width/height) are not suitable. In this paper, we assume the aspect ratio of the face is between 1/2 and 2.

Hair Area Search. The extracted facial region includes not only the face, but includes other skin color regions such as hands and desks that have similar colors. In order to discriminate the facial region from other regions, we use the hair region. If the skin color region has hair in the upper region, it is considered a face.

We check the intensity of the upper part of the skin color region. If the intensity is low, it is considered a fair candidate region. Then we search the entire area of the top skin color region, and if more than 20% of the area is black, the region is considered a head.

We show an example of the head extraction result in Fig. 4. In Fig. 4, the white region with small rectangle is the head region and gray region is a skin color region but not a head.



Input image Result image

Fig. 4. Result of skin color extraction and hair area search

2.2 Detecting a Raised Hand

We detect a hand that is in the area surrounding the face. A student will raise his/her arm to around the upper area of the face, thus, we set the region for detecting the hand in the following way.

$$S_{x1} = F_x - 5l/2$$
$$S_{x2} = F_x + 5l/2$$
$$S_{y1} = F_y - 2l/3$$
$$S_{y2} = F_y + 3l/2,$$

where S_{x1}, S_{x2}, S_{y1}, and S_{y2} are the minimum and maximum x and y values of the hand extraction rectangular area, F_x and F_y are the centers of gravity of the facial region, and l is the square root of the face area that is the length of the face. The large rectangle in Fig. 4 illustrates this region.

Next, we check whether the hand is moving in this area. We extract a moving object using frame subtraction and on erosion filter [5]. If the size of the extracted object is nearly the size of a hand and the object has skin color, we then extract this object as a raised hand.

The rectangular regions detecting the hands may overlap due to the seating arrangement of the students. If a hand is detected in an overlapped region, we must establish which person raised his hand. In this paper, we decide which person it is based on the distance between the hand and the head. We choose the person nearest to the moving hand.

3 Automatic Monitoring with a Video Camera

After the system detects a raised hand, the system must record the student's sign language. In this section, we describe the control of the video camera.

3.1 Pan-Tilt Control

In order to film a student in the center of the view, we control the direction of the camera. People usually sign below their face. Therefore, we need to control the camera direction to locate the face center at a level of 5/8 height.

3.2 Zoom Control

Other students understand a question not only through sign language but also by lip reading. Therefore, the system should zoom into the questioner so that others can easily see their lips. In this paper, we control the zoom to fill the visual angle with the sign language region. We use $S_{x2} - S_{x1}$ as the width of the sign language region.

Zooming Out During Speech. If a questioner speaks using large movements as they sign, it is likely that their gestures may go outside the edge of the visual field. In this situation, it would be difficult for other students to recognize what is being signed, therefore, the system should zoom out to fit the questioner's gestures within the visual field.

The system detects a moving hand using frame subtraction and skin color detection, and if the moving hand is located at the boundary of the visual field, the camera then zooms out. To avoid zooming out too far, the system does not zoom out more than five individual times.

Zooming in During the Speech. If the questioner's image size is too small, other students may find it hard to see his/her lips. In this case, the system zooms in further on the student. If the hand region is concentrated in the center of the visual field, the system again zooms into the student. The system will not zoom in more than five individual times.

If the zoom changes frequently, other students may find it hard to watch the video. Therefore, the system pauses for a given time between zoom operations.

4 Experimental Result

To evaluate our proposed system, we conducted an experiment in a real lecture for the hearing impaired. Figure 5 illustrates the arrangement of this experiment.

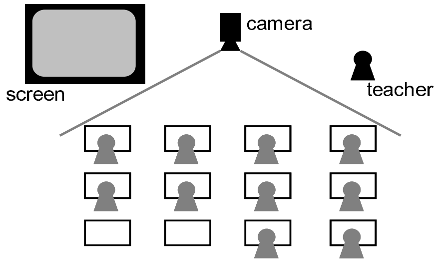


Fig. 5. Arrangement for the hand-raising detection experiment

Table 1. The number of raised hand detected

	Front	Middle	Back	Total
A	17	24	22	63
B	4	4	6	14
C	5	1	0	6
D	7	10	10	27
Total	33	39	38	110

Table 2. Raised hand detection rate

	Front (%)	Middle (%)	Back (%)	Total (%)
A	52	62	58	57
A+B	64	72	74	70
A+B+C	79	74	74	75

There were 10 subjects and they were seated randomly. All subjects were asked to raise their hand 11 times.

We evaluated the system by examining the camera image after zooming into the questioner. We use the following evaluation criteria.

A: Does the face and hand moving region lie inside the camera view?
B: Is a part of the region outside the camera view?
C: Is more than half the region outside the camera view?
D: Does the camera not zoom in?

We show the row-by-row results in Table 1 and the detection rates in Table 2. When a part of the region was out of the camera view (B), the system zoomed out from the questioner because the skin color region was located at the boundary of the camera view. After zooming out, the whole of the face and hand moving region was inside the camera view.

5 Discussion

We can see that 57 % of the raised hands were detected where the camera zoomed in correctly. When a part of the region was out of the camera view (B), the system zoomed out and recorded the entire region. Therefore, we consider the final detection rate to be 70%. In cases where more than half the region was out of the camera view (C), it was hard to recover the correct angle by zooming out. The system was able to recover by finding the head region again and changing the camera direction.

The reason why the camera sometimes did not zoom in to the questioner (D) is that the system failed to detect the raised hand or the system detected a different person. In some cases, the system could not detect the head position due to a failure of the skin color detection system. The system needs a more robust method for detecting skin color. The system sometimes detected a different person due to that person's low level of movement. When multiple raised hands were detected, the system chose the person nearest to the hands. We can improve the detection of a raised hand using a more suitable method based on the direction and amount of hand movements [6,7].

We also conducted questions and answers in the real lecture. When the system succeeded in zooming in on the questioner, other students can recognize what is being said by watching the questioner's sign language and his/her lips on the monitor. We can see the system is useful for sharing questions among all students.

6 Conclusion

In this paper, we proposed a system for detecting a questioner, monitoring his/her sign language, and displaying an image of the questioner on a monitor at the front of a lecture room for the hearing impaired. Since the questioner raises his/her hand when asking a question, we proposed a method of detecting a raised hand. We implemented this system and conducted an experiment in a real lecture for the hearing impaired. The experimental results show that the extraction rate of the raised hand was 70% and the system was useful for sharing questions among students.

Further work should be done to improve the detection rate of the raised hands. This method will improve the detection of the raised hands based on the direction and amount of hand movements. The system also needs a more robust technique for detecting a raised hand and for zooming on the questioner.

This research is partly supported by the Artificial Intelligence Research Promotion Foundation.

References

1. Quen-Zong Wu, Bor-Shen Jeng: Background subtraction based on logarithmic intensities. Pattern Recognition Letters **23** (2002) 1529–1536
2. Kwok-Wai Wong, Kin-Man Lam, Wan-Chi Siu: A robust scheme for live detection of human faces in color images. Signal Processing: Image Communication **18** (2003) 103-114

3. Hongxun Yao, Wen Gao: Face detection and location based on skin chrominance and lip chrominance transformation from color images. Pattern Recognition **34** (2001) 1555-1564
4. Chang-Woo PARK, Euntai Kim, Mingon Park: Human Face Detection via Characterized Convex Regional Relationship in Color Image. IEICE Trans. inf. & syst. **E86-D** 4 (2003) 759–762
5. Miltion Chen: Achieving Effective Floor Control with a Low-Bandwidth Gesture-Sensitive Videoconference System. Proc. ACM Multimedia (2002) 476–483
6. Ho-Sub Yoon, Jung Soh, YounglaeJ.Bae, Hyun Seung Yang: Hand gesture recognition using combined features of location, angle and velocity. Pattern Recognition **34** (2001) 1491–1501
7. Leonid V.Tsap: Gesture-Tracking in Real Time with Dynamic Regional Range Computation: Real-Time Imaging **8** (2002), 115-126

Efficient Generation of Large Amounts of Training Data for Sign Language Recognition: A Semi-automatic Tool

Ruiduo Yang[1], Sudeep Sarkar[1], Barbara Loeding[2], and Arthur Karshmer[2]

[1] University of South Florida, Tampa, FL, 33620, USA
[2] University of South Florida, Lakeland, FL, 33803, USA

Abstract. We have developed a video hand segmentation tool which can help with generating hands ground truth from sign language image sequences. This tool may greatly facilitate research in the area of sign language recognition. In this tool, we offer a semi automatic scheme to assist with the localization of hand pixels, which is important for the purpose of recognition. A candidate hand generator is applied by using the mean shift image segmentation algorithm and a greedy seeds growing algorithm. After a number of hand candidates is generated, the user can reduce the candidates by simple mouse clicks. The tool also provides a hand tracking function for faster processing and a face detection function for groundtruthing non manual signals. In addition, we provided a two-passes groundtruthing scheme unlike other tools that only does one-pass. Our first pass processing is automatic and does not need user interaction. The experiment results demonstrate that based on the first pass's result, one can groundtruth 10,000+ frames of sign language within 8 hours.

1 Introduction

With the growing capacity of computer vision technology, there is more interest being directed towards the automatic recognition of sign language based solely on image sequences or videos. We refer the reader to [1] for a review of the applications of related vision technologies. Briefly, if an accurate and reliable method can be developed to recognize and translate from sign sentences to spoken or written languages, this will help alleviate communication difficulties that Deaf people experience daily. This method can be incorporated into an assistive device carried by the Deaf person or available within the computer systems of stores, banks and other places where a Deaf person needs to communicate with non-signers.

In a vision-based sign language recognition system, generally a model will be trained on either a phoneme, sign or sentence level. This requires accurate feature extractions from the image frames. For example, one may need to extract the facial expression, hand position and hand shape parameters. However, the automatic hand segmentation is a challenging problem and unwanted noise will be generated for real world image sequences; hence the accuracy of the model

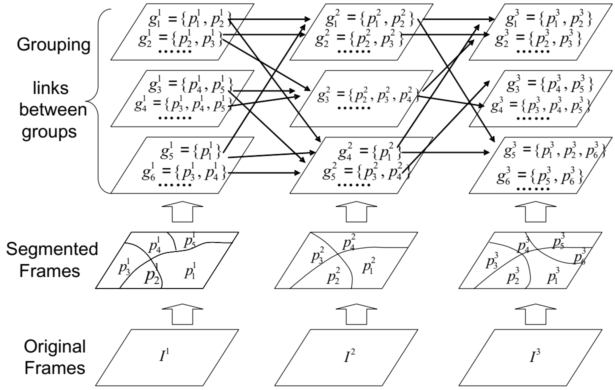
K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 635–642, 2006.

can be decreased. For example, an automatic scheme for segmenting the hand can be based on hand color model [2], stereo depth [3], colored gloves [4] or motion [5] information. Nevertheless, a hand color model is not sufficient for real world images where a large variation of lighting condition exists. Alternatively using depth image is slow and will fail when the hand moves to the same plane as the body. In addition, colored gloves might help during training but it still changes the appearance of hand and makes the signer feel unnatural. And finally, although using motion information is usually robust, it cannot accommodate subtle hand shape changes and a moving background. Due to these reasons, a hand groundtruthing step is often taken with an annotation tool before we start the training. Sign language training databases generally consist of many video shots and frames, which makes it very hard to groundtruth quickly using general image processing tool (e.g. Photoshop) frame by frame. Even with a vision based video annotation tool, this task could be labor intensive and time-consuming. For example, considering a small training database that has 100 signs, each sign has 25 training samples and the average length of one training sample is 50 frames. There are a total of 25x100x50=125000 frames. Suppose we were using a general image processing tool which can generate groundtruth for one frame in 5 min. The time cost to groundtruth the whole database will be 434 days.

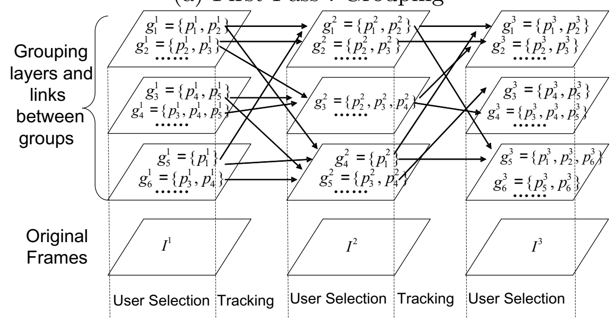
In this paper, we offer a semi automatic scheme to assist with the localization of hand pixels, which is important for the purpose of recognition. A candidate hand generator is applied by using the mean shift image segmentation algorithm and a greedy seeds growing algorithm. After a number of hand candidates is generated, the user can reduce the candidates by simple mouse clicks. In addition, we provided a two-passes groundtruthing scheme unlike other tool that only do one-pass. Our first pass processing is automatic and does not need user interaction.

We are aware of many other video annotation tools. However, most focus on scene segmentation or key frame detection (e.g. IBM EVA[6], ESP Game[7]). Some of them combine local feature extraction and temporal tracking together. For example, the VIPER annotation tool proposed by Pfund et.al [8] provides image segmentation, temporal segmentation and event annotation together; ViPer-GT proposed by Doermann et.al [9] can detect multiple objects and track them using bounding boxes automatically, Marcotegui et.al proposed VOGUE [10], where a number of image and video segmentation techniques are incorporated for object annotation purpose. All of these tools are standalone applications providing semi automatic groundtruthing function with user-friendly interface.

Our annotation tool is a side-product of our vision-based American Sign Langugae(ASL) recognition system. It also provides a semi-automatic scheme for efficient ground-truthing. However, its main purpose is to segment the hand pixels frame by frame. Instead of using the general segmentation and tracking algorithm as in the existing tools, we advocate a candidate hand generator approach which is more reliable during hand shape change and hand crossing face situation. Unlike the existing tools where only one-pass is conducted, we offer a two-pass scheme for faster processing, where the first pass generates candidate hands automatically.



(a) First Pass : Grouping



(b) Second Pass: Selection

Fig. 1. Overview of the two-pass approach

Fig. 1 illustrates our two-pass scheme. In Fig. 1(a), ASL video frames are first segmented into seed primitives; these primitives are grouped by a grouping engine to generate overlapped candidate hand groups, where each group may consist of one or more primitives. This step is automatic and no user interaction is involved. In Fig. 1(b), where the second pass is taken, the grouped results will be loaded back for user's examination. Note the number of generated groups could be huge, hence we allow the user to mouse click the hand region to reduce the number of candidates. At the same time a tracking method is also incorporated among adjacent frames to improve efficiency.

The paper is organized as follows: in section 2 we will discuss the algorithm used in the first-pass, Section 3 describes the Graphic User Interface(GUI), Section 4 shows us the experimental results and we conclude with Section 5.

2 First-Pass: Segmentation and Grouping

The first-pass annotation can be done completely on the background since there is no user interaction involved. The goal of this pass is to generate the hand candidates and store them as a file to be reloaded in the second pass. For existing annotation tools, these two passes are generally working at the same time, forcing the user to wait for the segmenting and tracking result. With a complex scene, the segmentation itself can be very time consuming. (e.g. a segmentation of one single frame with a highly textured scene with mean shift method [11] could take over 10 seconds). On the other hand, many tracking techniques take an iterative approach which could also be slow. Tracking methods also suffer from the requirement of good initialization, which must be taken for each sign in the training dataset separately.

For our case, since we tend to generate as many candidates as possible to accommodate with the fast changing appearance of the hand while signing, the processing of one single frame will even be slower. For example, with a complex scene there will be over 500 candidate hands generated, which result in a running time for one frame to exceed 30 seconds.

Fortunately, unlike other tracking or hand segmentation schemes, our approach does not require perfect results or user modification when the candidate hands are being detected. Hence this step can be done solely in the background.

2.1 Grouping of Primitives

We adopt a greedy approach to form the groups. Let us denote the low-level region patches in the k-th image frame by $S_k = \{p_1^k, \cdots, p_{N_k}^k\}$. From this initial set of primitives, we select a subset of primitives that are likely to come from the hand, based on the size of the patch. These are our seed patches. Given some knowledge of the approximate size of the hands in the sequence, we can eliminate large, non-homogeneous region patches from further consideration. We use a list L to store the possible groups. This list is initialized by choosing each selected primitive to be a singleton group. These groups would be merged to form a larger conglomerate.

The grouping process starts by picking the first group in L, denoted here by p, and searching for its neighbors $\{N_p^i\}$. Each neighbor N_p^i is considered for grouping with p to generate a tentative larger grouping. We select the best local grouping, and denote it as g. The group g is further tested to see if it can possibly represent a hand. If the group g passes this test, it is inserted into the final candidate group list, C, otherwise it is inserted at the end of the list L, to be considered for further grouping.

2.2 Associating Groups Across Frames

We denote the jth group detected in tth frame as G_t^j. The groups detected in each frame are associated with those detected in previous frames to result in a linked sequence of groups spanning all the frames. We define the predecessors set of each element in each groups set as

$$Pre(G_t^j) = [G_{t-1}^{j_1}, \cdots, G_{t-1}^{j_n}], t \geq 2, 1 \leq j_k \leq c_{t-1}, \quad (1)$$

where $G_{t-1}^{j_k}$ is one possible predecessor of G_t^j. The predecessor relationship between the groups from different time instants is based on feature similarity. It captures how likely the groups are from the same underlying cause in the image. We will use the best successors as the tracking result.

3 Second Pass: Graphic User Interface

After the candidate groups and their links are generated, user interaction is needed to select the best group, with a default group given at each frame by the tracking method. We provide functionality that specifically works well for sign sentences. These functions are built upon the candidate generator discussed in section 2, a simple tracking technique that work with the links between the candidate groups, a face detector for non-manual information analysis, a glossing tool and various elements that facilitate the hand groundtruthing.

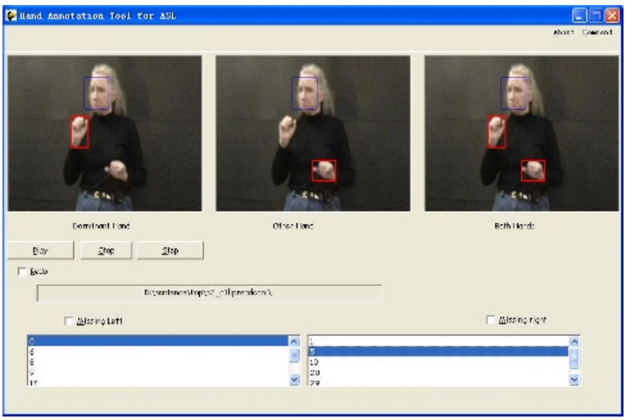


Fig. 2. The Graphic User Interface

The application is coded under the Microsoft Visual Studio Environment, using MFC class, OpenCV libraries and related windows APIs. Fig. 2 illustrate us the GUI. Important functions related to sign language are supported as:

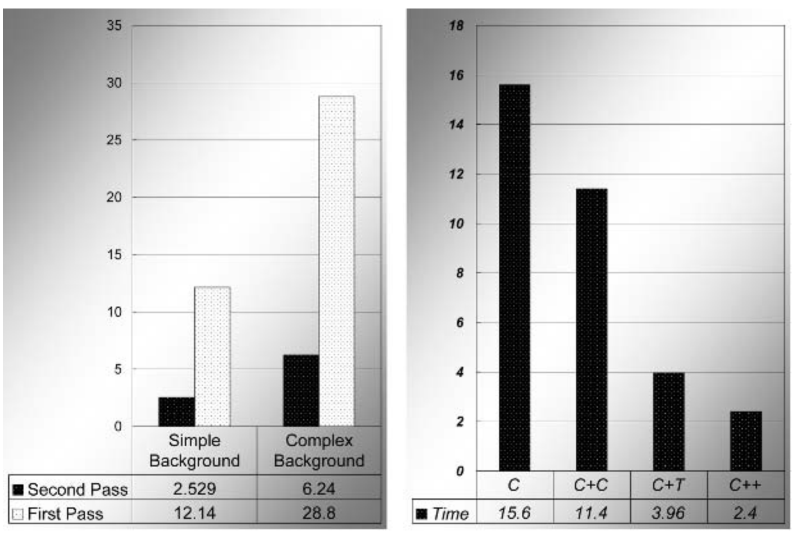
1. 3 views of current frame: The first is the dominant hand view, the second is the non-dominant hand view, the third is the view for both.
2. Click to select: One can click on the hand area to reduce the groups to be shown. The shown groups must include the point where the mouse is clicked.
3. Missing hand checkbox: One can choose the current hand as missing if the hand is out of the scene.
4. 2 Hand listboxes: The two listboxes below show the list of the candidate dominant hand and candidate non-dominant hands. The list is ranked by their boundary smoothness and the tracking result. The highlighted item is the candidate hand that is being selected.
5. Face detector: automatic face detector, shown as the blue bounding box.

6. Play,Stop,Step Button: Pressing the play button will automatically track all the hands and save the result. Stop Button will stop the tracking. Step button is to allow the application to track one frame and wait for the user's response, which is most often used.
7. Redo Checkbox: Redo Checkbox allows one to re-detect the current sequence.

4 Experiment Result

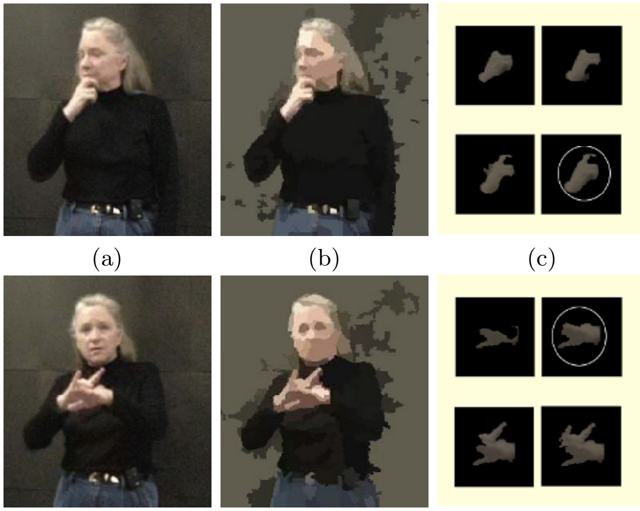
We tested our annotation tool with 2 datasets with different parameters settings. Both of them consist of ASL sign sentences. The first dataset has a simple background with 10675 frames and a resolution of 460x290. On average there are 100 candidate hands generated per frame for this dataset, yet it took us less than 8 hours to finish groundtruthing both hands. The second dataset has a complex background with 640x480 resolution. There are 500 candidate hands generated for each frame. We took 500 frames and established the groundtruth within 1 hour. Note the time we refer to here is the user interaction time, that is, the time of the second pass.

In Fig. 3(a), we show the time taken over the 2 datasets of the two passes. We use a P4 2.4G CPU with 4G memory. The number shown is the time taken per frame in seconds. Our first pass takes relatively longer since we incorporated automatic segmentation, face detection, and the greedy seeds growing algorithm. However it's completed offline. The second pass is done by reloading the candidate result, taking much shorter time.



(a) (b)

Fig. 3. (a) Time taken of first pass and second pass (in seconds). (b) Time taken with different method sets. (in seconds).



(a) (b) (c) (d) (e) (f)

Fig. 4. (a) Original frame 1. (b) Segmented frame 1 (c) List of candidate groups in frame 1. (d) Original frame 2. (e) Segmented frame 2 (f) List of candidate groups in in frame 2.

On the other hand, Fig. 3(b) shows us the time taken for different method sets of the tool. We choose 500 frames from the simple background set to do the experiment. Here we refer to "C" as the method that only uses the generated result, "C+C" refers to using the click-to-select method to reduce the candidate set, "C+T" refers to using candidate with tracking method, "C++" refers to the method of using candidate and both click-to-select and tracking method. Tracking contributes a lot because it exploits the temporal relationship. The click-to-select method does help when tracking failed. For example, when a large motion happens, drastic hand shape changes and occlusion happens. At the same time, we did an experiment, using wand tools in a image editing software to groundtruth 10 frames randomly chosen from the dataset. The average time taken for each frame exceeds 3 min.

Fig. 4 shows us some visual results of the generated candidate hands. We can see our method can discern the handshape even when there is occlusion and overlaps. In particular, Fig. 4 (c) shows us the result where face crossed hand, and Fig. 4 (f) shows us where the two hands crossed each other.

5 Conclusion

In this paper, we describe our semi-automatic tool for annotating sign language video sequences. The objective of this tool is to generate large amount of data with labeled hand pixels that can provide reliable training process. The tool uses segmentation, tracking and offline processing method to considerably save the time to get the reliable hand data. For future work, an assisting edit tool can be added to secure the situation when there are merges among hand pixels and non-hand pixels where segmentation and grouping fail.

Acknowledgment

This work was supported in part by the National Science Foundation under grant IIS 0312993. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect those of the National Science Foundation.

References

1. B. L. Loeding, S. Sarkar, A. Parashar and A. I. Karshmer.: Progress in Automated Computer Recognition of Sign Language., Lecture Notes in Computer Science, **3118**(2004), 1079 - 1087
2. T. Starner, and J. Weaver and A. P. Pentland.: Real-time American Sign Language Recognition using Desk and Wearable Computer Based Video, IEEE Transactions on Pattern Analysis and Machine Intelligence, **12** (1998) 1371–1375.
3. L. V. Tsap and M. C. Shin.: Dynamic Disparity Adjustment and Histogram-Based Filtering of Range Data for Fast 3-D Hand Tracking, Journal of Digital Signal Processing, **14** (2004) 550–565.
4. A. Just and S. Marcel.: Two-Handed Gesture Recognition, IDIAP Research Institute CH-1920, 2005, Martigny, Switzerland
5. M.H. Yang., N.Ahuja. and M. Tabb.: Extraction of 2D Motion Trajectories and its Application to Hand Gesture Recognition., IEEE Transactions on Pattern Analysis and Machine Intelligence, **24** (2002) 168–185.
6. T. Volkmer and J. R. Smith and A. Natsev.: A Web-based System for Collaborative Annotation of Large Image and Video Collections: An Evaluation and User Study, MULTIMEDIA '05: Proceedings of the 13th annual ACM international conference on Multimedia, (2005) 892–901.
7. L. V. Ahn and L. Dabbish.: Labeling Images with a Computer Game. In Conference on Human Factors in Computing Systems (CHI) (2004) 319–326.
8. T. Pfund and S. M. Maillet. A Dynamic Multimedia Annotation tool. In G. Beretta and R. Schettini, editors, In Proceedings of SPIE Photonics West, Electronic Imaging 2002, Internet Imaging III, **4672** (2002) 216–224.
9. D. Doermann and D. Mihalcik. Tools and Techniques for Video Performance Evaluation. In Proceedings of the IEEE International Conference on Pattern Recognistion (ICPR'00), **4** (2000) 167–170.
10. B. Marcotegui, P. Correia,A Video Object Generation Tool Allowing Friendly User Interaction.,International Conference on Image Processing, **4** (1999) 391–395.
11. D. Comanicu and P. Meer.: Mean shift: A Robust Approach Toward Feature Space Analysis., IEEE Transactions on Pattern Analysis and Machine Intelligence, **24** (2002) 603–619.

Composition Corrector – A Computer-Based Tool to Support Professional Deaf and Hard-of-Hearing Writers

Martha Birnbaum[1], Corine Bickley[2], and Brianne Burger[3]

[1] Martha Birnbaum Consulting, 7 Hancock Place, Cambridge, MA 02139
martha@marthabirnbaum.com
[2] Department of Hearing, Speech and Language Sciences, Gallaudet University, Washington, DC 20002
corine.bickley@gallaudet.edu
[3] Department of Administration and Supervision, Gallaudet University, Washington, DC 20002
brianne.burger@gallaudet.edu

Abstract. *Composition Corrector* (*CC*) is an online, browser-based tool that detects and corrects grammatical errors in written English composition. CC is aimed at deaf high school and college students who are fluent in a language such as American Sign Language but are not necessarily fluent in written English. The student types an English sentence into the browser window and receives an immediate analysis of grammatical errors and the corrections in standard English. The writing skills of deaf students are reported to be significantly poorer than those of their hearing peers, with attendant educational, professional, and income disparities with the hearing population. *CC* is conceived as a tool to present students to best advantage in an academic or professional setting. It acknowledges a social as well as educational need in that strong English composition skills are companions to independence, quality of life, and professional development.

1 Introduction

Composition Corrector (CC) is an online, browser-based tool that detects and corrects grammatical errors in written English composition of deaf students whose first language is ASL (American Sign Language) or other signing systems. CC is aimed at a critical and underserved population – profoundly deaf students who communicate fluently in sign. The writing skills of deaf students are consistently reported to be significantly poorer than those of their hearing peers, and it is not unusual for a deaf student to graduate from high school with the written communication skills of 9-10 year olds [1,6]. A serious consequence of this gap in writing competency for deaf individuals is a limitation in educational and professional arenas, where writing difficulties are a significant impediment to higher education and access to many professions [4]. Lack of professional advancement is accompanied by salary disparity, almost guaranteeing that deaf workers with weak English skills will be chronically underpaid.

CC leverages three salient facts: the ubiquity of computers, the Deaf community's enthusiastic and skillful acceptance of useful technical innovations, and the desire of

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 643 – 650, 2006.

deaf individuals to have increased access to educational and professional opportunities as independent professionals. With CC, the student types an English sentence into the browser window and receives an immediate analysis of grammatical errors and the corrections in standard English. CC uses natural language understanding algorithms to process errors and is trained on a large corpus of deaf students' compositions.

CC is conceived as a tool to give students independence in crafting grammatically correct essays, and to present them to best advantage in an academic or professional setting. It acknowledges a social as well as educational need in that strong English composition skills are companions to independence, quality of life, and professional development.

2 Grammar-Correction Tools

Prior computer-based approaches to improving the writing skills of deaf students such as the ICICLE system [5] have used tutorial dialogues rather than direct analysis and correction of errors. The grammatical correction tool most familiar to everyone with a Windows-based computer is Microsoft's Grammar Checker, a feature of Microsoft Word. We selected 494 sentences at random from our corpus of deaf students' sentences and ran them through both Word Grammar Checker and *Composition Corrector*. The results show that *CC* outperformed Grammar Checker in all tests: for every error category, Grammar Checker found only half or fewer of the errors that *CC* found. Further, Grammar Checker does not usually find more than one error in a sentence, whereas *CC* finds almost all multiple errors. *CC* was developed specifically to detect and correct the errors found most commonly in English sentences written by deaf students; Grammar Checker is aimed at a more general user group. There are currently no other commercial products that accept typed input and detect and correct errors *in real time*.

3 Research Methodology

3.1 Data Collection

In order to develop the algorithms that are the basis of parsing and grammatical error detection and correction, we needed a large amount of written data representative of our eventual users. Research in journals, periodicals, and on the Internet revealed that there is no publicly available database of English text written by deaf students. Accordingly we designed a data collection method that would provide us with a large number of sentences and the ability to collect additional data whenever we needed it.

We designed and built a computer-based data collection application that consists of a demographic questionnaire, informed consent form, 21 pages of writing tasks, and a page for the student to evaluate his/her experience as a participant. Students were paid to complete the 1-hour 15-minute task, and were supervised by a research assistant who is deaf and natively bilingual in ASL and English. The browser-based application and an hsqldb database to hold the input sentences are housed on a commercial web hosting server.

The data collection application comprises three types of writing activity in response to pictures. In Task 1, students are asked to construct sentences about the pictures from sets of 3 or 4 words. The words in the sets are content words and the form of the sentence is left to the student's choice. In task 2, students fill in empty slots in sentence frames relating to the picture in order to make a grammatical English sentence. In task 3, students write a paragraph about the captioned picture(s).

Using the online data collection tool, we collected *8233* sentences and an additional *1239* sentences in a free essay task [2]. The complete corpus represents a unique body of data in the textual language arena and is, we believe, the only substantial corpus of the writing of deaf students who sign.

3.2 Analysis of Patterns

The basis of *CC* is the observation that the English writing of deaf students contains the syntactic categories of standard English but has *different yet systematic* ways of implementing these categories. *CC* treats these systematic patterns in the writing of deaf individuals as realizations of patterns that have a "non-standard" label. It is probable that these different yet systematic patterns are based on whatever signed language a student uses and the conditions imposed on communicating information in a sign language. Parsing of deaf students' written English then becomes a task of creating rules to recognize deviations from standard patterns. We refer to these non-standard patterns as "errors" because any non-standard pattern in a grammar correction system would be so designated. We are guided in the definition of standard grammar and grammatical categories and their relation to ASL by Bordman and Womeldorf, *The Gallaudet Writer's Handbook, 1999.*

The algorithms underlying CC are based on an analysis of the written (non-standard) errors of deaf students contained in the database described in 3.1. These algorithms use natural language understanding technology to parse a sentence, detect error patterns, and map the errors to correct standard English patterns [3].

3.3 Error Taxonomy

Table 1 shows the broad types of errors observed in the database through hand analysis.

Table 1. Error categories and examples

Description	**Example**
Verbal errors	
subject-verb agreement	The boy go to school
verb conjugation error	He is plays the piano.
auxiliary-verb error	He may wants to go.
Infinitive errors	
missing "to" infinitive marker	They want go home.
infinitive verb form error	She wants to sings
Nominal errors	
qualifier-noun agreement	Each books are here.
qualifier-prep phrase agreement	All of the book is here.
possessive-noun agreement	The man saw she book.

These types comprise verbal errors involving subject-verb agreement, verb conjugation errors, and auxiliary-verb agreement; errors involving infinitives; and nominal errors, including qualifier-noun agreement.

3.4 CC Engine

Input words are tagged using an innovative multi-level tagging algorithm and inferential rules based on context. A set of parsing rules with ordered application analyzes strings of tagged words, and a set of clause modules establishes the hierarchic relationships among main and embedded clauses. Current work is focused on expanding clause modules and developing heuristic algorithms to learn and incorporate new grammatical patterns.

Table 2. Bi-level part of speech tagging

Tags	**Word1**	**Word2**	**Word3**	**Word4**
	The	*dog*	*want*	*run*
deepTag	ArtD_x	Noun_s	Vb-4xrx	Vb-4xrx
posTag	QUL	N	V	V

Table 3. Composition Corrector error detection modules

Error	**Description**	**Example**	**Correction**
SVA	subject-verb agreement	The boy go to school.	The boy goes to school.
VCJ	verb conjugation error		
	be-verb agreement	He is plays the piano.	no correction performed
	have-verb agreement	He has go.	no correction performed
	do-verb agreement	He do go.	no correction performed
INF	infinitive error	She wants to singing.	no correction performed
MST	missing “to” infinitive marker	They want go home.	They want to go home.
QNA	qualifier-noun agreement	Six man came home.	Six men came home.
QPA	qualifier-prepositional phrase agreement	All of the book are here.	All of the books are here.
AVE	auxiliary-verb error	He may wants to go.	no correction performed
PNA	possessive pronoun agreement	The man saw she book.	The man saw her book.
NGE	no grammatical errors	The dog reached the roof.	no correction performed

Besides the parsing engine, CC comprises a 4000-word dictionary, a database, an easy-to-use online graphical interface, and a statistical analysis platform that calculates word and phrasal probabilities.

Table 2 shows **"*bi-level tagging*"©,** a CC method in which each word receives a "deepTag" containing as much detailed grammatical information as possible, and a "posTag" (part-of-speech tag). The output of the tagger is a set of consecutive input words, each with a deepTag and a posTag label. In the example in Table 2, the sentence "The dog want run." from the database received the following tags:

posTags are higher-level grammatical categories such as N(oun), V(erb), QUL (Qualifier). deepTags indicates detailed information such as verbal person, number, tense, and gender, and nominal number. The strength of bi-level tagging is the power it gives the parser to distinguish grammatical from ungrammatical combinations of deepTags, or deepTags + posTags, and to use that information "on the fly" to correct the tagged sentence while the tagger is running.

Parsing rules use these tags in parsing strategies to detect and correct errors. Beginning with the broad error categories of Table 1, and using the tagger and a training subset of the database sentences, we developed rules for automatically detecting 7 error categories, and also a category NGE (no grammatical errors). The rules are listed in Table 3. Column 1 shows the error category acronym, column 3 gives an example, and column 4 shows the correction that the CC parser outputs.

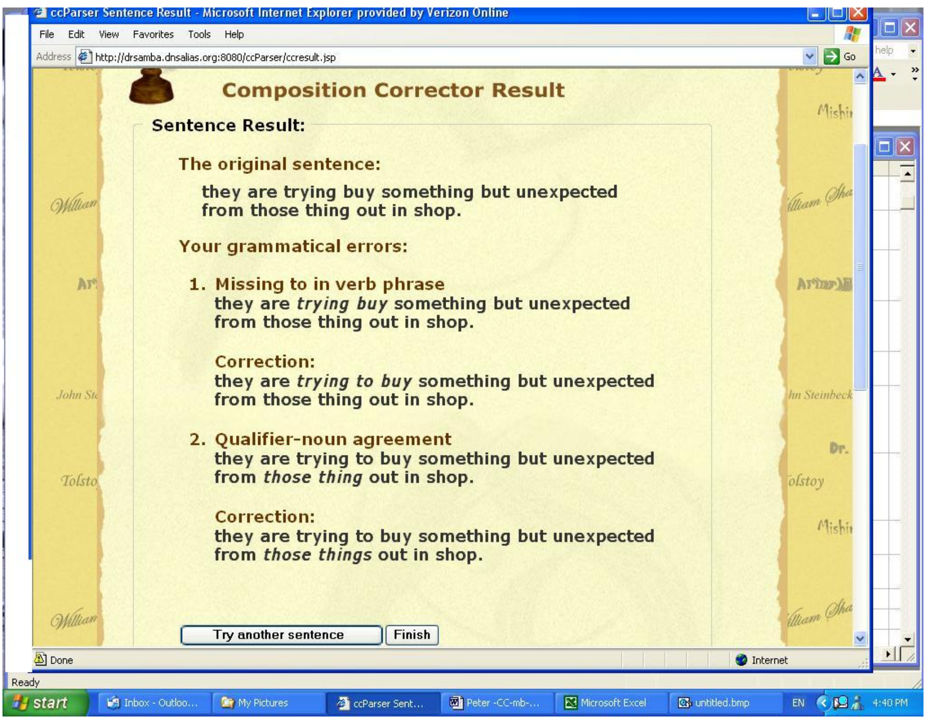


Fig. 1. A *CC* results page showing input sentence, errors, and corrections

Note that some errors are not corrected, but merely indicated to the writer to be ungrammatical. The parser is conservative in that it does not offer a correction when the writer's original intent is ambiguous or unclear to the program. In these cases, the parser simply indicates the error by marking the words involved.

3.5 User Interface

Figure 1 is a screen capture of the operation of CC on a complex sentence from the corpus: "*They are trying buy something but unexpected from those thing out in shop*". This sentence contains multiple errors, including instances in which the intent of the writer is not clear and CC cannot guess that intent (nor could a human). CC does correctly detect and correct two clear errors of a missing "to" and non-agreement of qualifier and noun. All correction steps are displayed to the user.

4 Results

In the testing of the prototype system, CC had a 91.6% correct average score over 7 types of errors for 2704 sentences chosen at random from the database. Sentences were from the first data collection task, in which students were asked to construct sentences using three or 4 given words related to a picture. Each of the sentences was hand-analyzed as exhibiting one or more errors or as exhibiting no grammatical errors. The sentences were input into CC in batch mode. Table 4 compares the results of hand analysis and *CC* analysis of the sentences by error type. *CC* performance is shown in terms of *precision*, a measure of how often the system correctly found an error.

Table 4. Task 1 sentences: total number of sentences analyzed- 2704

error type	errors found by *CC*	false positives	errors found by hand	precision
SVA	402	19	382	95.3%
QNA	36	2	34	94.4%
VCJ	24	1	23	95.8%
MST	16	0	16	100.0%
QPA	16	5	11	68.8%
INF	9	0	9	100.0%
AVE	8	1	7	87.5%
average performance over 7 error types				91.6%

CC found no errors in 2120 sentences. Hand analysis of those sentences confirmed the automated *CC* analysis in nearly 100% of the cases; that is, hand analysis deemed 1 of the *CC* "errors" to be a false positive. This result confirms the basic soundness of the *CC* parsing algorithms to distinguish between errors and non-errors.

For the largest error category, SVA (subject-verb agreement), *CC* found 402 errors, 19 of which were determined to be false positives by hand-analysis. *CC* performance for SVA was thus 95.3% accuracy. The performance of *CC* for the error categories QNA (qualifier-noun agreement), VCJ (verb conjugation error), MST (missing "to" infinitive), and INF (infinitive phrase error) also show excellent performance - between 94% and 100% agreement with hand-analysis. QNA false positives occurred mainly with compound noun phrases, which were not within the scope of the initial CC system, and are currently being addressed. An example is: *Both young bear polar have always fun in anywhere*. *CC* corrected this to: *Both young bears polar have always fun in anywhere.*

QPA errors are those in which a noun in a prepositional phrase does not agree with its antecedent qualifier. An example from the corpus is: *One of the cute cub climbs the rowboat*, corrected to "*cubs*". False positives again involved compound noun phrases: *One of the bear cub climbs the rowboat*, falsely corrected to *One of the bears cub climbs the rowboat.*

AVE errors are auxiliary verb errors. An example of an AVE error is *The girl won't looked at the boy.* INF refers to an incorrect infinitive construction. An example is *The dog wants to reaching the roof.* These two categories accounted for the fewest number of errors in the corpus and are currently being addressed.

5 Future Development

In the coming development cycle, CC will be beta-tested at three American schools with sizable deaf populations: Gallaudet University in Washington DC, the CAEBER Institute of the New Mexico School for the Deaf in New Mexico, and Northern Essex Community College in Massachusetts. It will also be tested at Advocates, Inc. a non-profit service human services agency in Framingham, Ma with several deaf staff members.

Our immediate plans are to develop *CC* as a tool for deaf students in high school and college. A future version of *CC* will be for younger deaf children who are learning written English. *CC* will be easy to use because almost all children are now using a computer at an early age and are quite adept at game playing and Web surfing. *CC* will be accessed either through a CD that a parent or teacher purchases or through online subscription. It can be implemented on a home computer or on a language or writing laboratory computer under the auspices of a school or university.

In future, *Composition Corrector* will be adapted as a pedagogical tool for ESL students in university English courses, and as a tool for document correction for businesses and government functions. Currently the largest foreign language markets are Spanish, Chinese, and Japanese.

References

1. Albertini, J. A., Schley, S.: Writing: Characteristics, Instruction, and Assessment. In: Marschark, M., Spencer, P. E. (eds.): Deaf Studies, Language, and Education. Oxford University Press, Oxford (2003)
2. Bickley, C., Birnbaum, M., Moseley, M.J., Burger, B.: Patterns in English Sentences Written by Deaf College Students. American Speech Language and Hearing Association. San Diego (2005)

3. Bickley, C., Birnbaum, M.: Composition Corrector: A Computer-based Tool to Improve English Compositions of Deaf Students. RESNA, Atlanta (2005)
4. Marschark, M.: Psychological Development of Deaf Children. Oxford University Press, New York (1993)
5. Michaud, L. N., McCoy, K.F.: Modeling User Language Proficiency in Writing Tutor for Deaf Learners of English. Proceedings of the Symposium on Computer-Mediated Language Assessment and Evaluation in Natural Language Processing, an ACL-IALL Workshop. College Park, MD (1999) 47-54
6. Paul, P. V.: Literacy and Deafness: The Development of Reading, Writing, and Literate Thought. Allyn and Bacon, Boston (1998)

VoxAid 2006: Telephone Communication for Hearing and/or Vocally Impaired People

Bálint Tóth and Géza Németh

Department of Telecommunications and Media Informatics (TMIT),
Budapest University of Technology and Economics (BME),
Magyar Tudósok körútja 2., 1117 Budapest, Hungary
{toth.b, nemeth}@tmit.bme.hu

Abstract. Speech and/or hearing impaired people have difficulties with voice communication. In case of face-to-face conversation they can find a common communication channel (e.g. sign language, paper, etc.), but without an appropriate system they are unable to talk over the phone. The goal of the present study is to introduce the design and development steps of a system for vocally and/or hearing impaired people, which helps them to communicate via telephone with any person. Speech output is realized by text-to-speech (TTS) technology and speech input is provided by automatic speech recognition (ASR). The visual and the speech user interfaces enable users on both side of the phone line (a speech and hearing impaired person at one end, a non-speech-and-hearing-disabled person at the other end) to communicate.

1 Introduction

In the telecommunication world new technologies have been developed and released so rapidly as never before, which changes everyday life. Unfortunately the demands of disabled users are rarely considered during the development phase. Present and future technologies of telecommunication enable impaired persons to achieve a better standard of living. However if their special demands are not considered and if they are not involved in the development phase they can end up in a very hard situation.

The number of severely speech impaired persons is about two million, and the number of severely hearing impaired people is around six million in the European Union [1]. One can loose his/her ability to speak and/or to hear because of illness (e.g. stroke), physical impact (e.g. accident, surgery) or s/he might be already born as an impaired person.

Impaired hearing and speech causes many difficulties for the disabled person and also a lot of efforts for her/his environment (e.g. family). The lack of the above-mentioned abilities may also result in social and psychological problems.

Therefore a communication aid, which enables conversations over the phone for vocally and hearing impaired persons, can help them not only in everyday life, but it also helps them to approach the standard of living of non-disabled persons.

The aim of the present study is to investigate the possible user groups, to analyze previous assistive methods and devices for hearing and speech impaired persons, to

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 651 – 658, 2006.

give evidence for the necessity of a new system, and to introduce the design and development problems of a new communication aid.

2 Problem Statement

In the current chapter we investigate possible user groups and previous assistive technology solutions; furthermore we summarize the reasons why a new system is necessary.

User Groups

There are three possible user groups of the system: the hearing impaired; the speech impaired; and the both hearing and speech impaired people. Let us have a short investigation of these groups.

Hearing impairment can develop both at short and at wide frequency ranges. In the case of the phone conversation the 300..3400 Hz range is critical, as most phone systems transmit only this section of the frequency domain. In case of hearing impairments speech impairments often develop as well, as the speaker doesn't have a feedback of her/his voice. Dialing doesn't mean a problem for hearing impaired people, but they cannot hear the feedback (e.g. busy signal) and alert (e.g. ringing) signals. These problems can be solved with visual signals on the device. After a call is established the understanding of speech is a major problem. Persons with moderate hearing impairment can use their hearing aid with special inductive coupling [2], but this doesn't give a solution for severely hearing impaired people.

Speech impairment may affect the whole or a part of the speech generating process. Cellular phones and the Internet help a lot for speech impaired people, but voice based phone conversations still pose a problem for them. Impaired speech is often caused by moderate/severe mental disabilities. There are some communication aids [3] for speech impaired persons, which will be briefly introduced in the present paper.

Compared to the two previous groups the both vocally and hearing impaired people are in the most difficult situation. Their number is much smaller, but they have the greatest need of assistive technology solutions besides the groups mentioned above. Currently if one can afford it in Hungary s/he may have a sign language translator, but because of financial reasons it is affordable only for special, important events. There are halfway means for them (e.g. using computer chat), but without a communication aid they are unable to use the phone.

Present Solutions

There are quite a few assistive technology solutions for speech and/or hearing impaired people. The most important communication aids for speech impaired persons (e.g. BlissVox [4], Multi-Talk [4], Voxaid [5], SayIt for PDA [3] and Smartphone) have been investigated in [3]. For speech-and-hearing impaired people, according to the authors' knowledge, there is no automated solution for voice communication through the phone line, although there are systems that help impaired persons to communicate with others.

Text-phones enable chat based communication over the POTS (Plain Old Telephone System). A small keyboard and an LCD display with 2-3 lines are built into these devices, like the Uniphone 1150[1]. The disadvantage of text-phones is that on both sides of the phone line a compatible device is required. As the Internet became widespread, online technologies (e.g. instant messengers like ICQ[2] and Skype[3]) made text based communication easier, although the same disadvantage arises; on both side a computer with Internet connection is necessary.

The experiments in [6] showed, that the Finnish language is suitable for speaker dependent speech recognition on the phonemic level (the error rate is about 5%). In addition, according to the authors of the paper the 10 to 20 percentage of the phoneme errors can be tolerated with satisfying - word, sentence and dialog sentence - comprehension. These features are exploited in Finnish communication aids for deaf people.

For Swedish the acceptable error rates may be even smaller, as the orthography is less phonetic. It was reported in [7] that there is an about 26 percentage increment in tracking score and 17 percentage decrement of blocking-ratio in case of speech impaired persons when they used phoneme recognition as a hearing aid in telephone communication.

There are non-automated solutions for speech and hearing impaired persons to communicate with non-disabled persons through the telephone line in several countries of the EU (e.g. the Typetalk service in the UK). In Hungary the biggest Hungarian phone company offers the so called ‘Jelmondó’ (Sign-talker) service which was introduced in 2001. Practically it is a text-based communication over a ‘speech-gateway’. The text communication is realized with a computer and a modem or with a special device, and the ‘speech-gateway’ is a person, who ‘converts’ text-to-speech and speech-to-text. The service is free for the users although it costs money for the phone company (and the company cannot expect extra gain from the service) furthermore it is not suitable for confidential and personal conversations.

As it can be seen, there are existing systems for speech and/or hearing impaired people, but none of them offers adequate telephone features. Our aim was to create an automated communication aid for making and receiving phone calls, which can be used by the impaired person without any help. Furthermore it is also important, that the system doesn’t require any device on the other end of the phone line.

3 System Design

Generally the system should be able to convert text-to-speech (TTS), speech-to-text (STT) and to make and receive phone calls. Additionally the intuitive user interface is also important on both sides of the phone line; the speech impaired person should be able to use the system as fast as possible and on the other side of the line people should know that s/he is speaking with a speech and/or vocally impaired person, and not with e.g. a dialog system. The current system is called VoxAid 2006, named after our first communication aid, which was developed in 1990 [5].

---

[1] `http://www.teletec.co.uk/minicoms/uniphone.php`
[2] `http://www.icq.com/`
[3] `http://www.skype.com/`

Architecture

The system is implemented on Windows XP compatible computers, and an additional ISDN modem is required. Consequently the system on the user side operates on ISDN phone lines; the other side of the line can be of any kind. The ProfiVox TTS system is included as a dynamically linked library; the managed code of VoxAid 2006 operates perfectly with the native TTS code. Furthermore VoxAid 2006 communicates asynchronously with the ISDN interface (called CallCentre) via TCP/IP sockets with predefined messages in a handshaking way (Fig. 1).

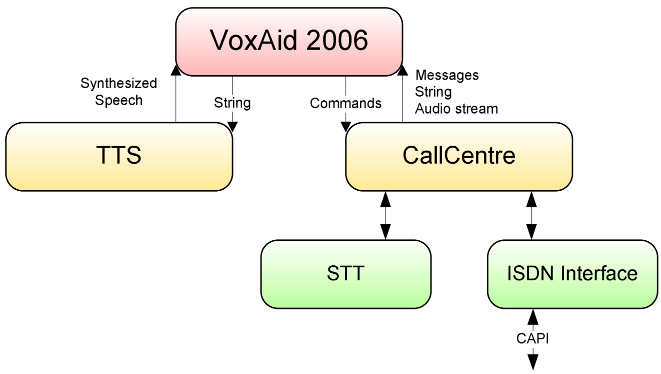


Fig. 1. The architecture of VoxAid 2006

The messages are transferred in byte arrays; one part of them is sent / received without parameters (e.g. playabort, playaborted, dtmfrec, etc.), others have parameters (e.g. play 'filename', recognized 'word', etc.). The HMM - based, speaker independent, Hungarian, telephone speech recognizer with fixed vocabulary is closely coupled to CallCentre. The ProfiVox TTS, the CallCentre ISDN interface and the speech recognizer are existing technologies, which have been developed in our department. The voice coming from the other side of the phone line can be turned on/off in the application, as speech impaired people without hearing impairments may use it. The voice is transferred as an audio stream through the ISDN interface with TCP/IP sockets.

During a call VoxAid 2006 sets CallCentre to speech and DTMF recognition mode. If the voice activity detection (VAD) alerts, DTMF recognition is aborted, and as soon as the speaker stops talking CallCentre sends the recognized unit (word or sentence) to VoxAid 2006, which displays it. If a DTMF signal is detected (for help or special settings, see below) than the ASR is stopped, and the code of the recognized DTMF signal is sent to VoxAid 2006.

Human-Computer Interaction (HCI)

The two main HCI parts of the application are the visual and the telephone user interfaces. As regards the architecture, the visual user interface consists of two main sections: speech generation and recognition (Fig. 2.).

Users have the possibility to define the layout of the screen. This feature helps people even with moderate mental disabilities to use the application, as it is possible, that one can perform some tasks only if the task can be done in a particular location of the screen.

Speech generation can be done in three modes: free text (the user can enter any text), fixed text (the user can select from previously stored, categorized sentences) and partly fixed text (similar to the fixed text mode, but the user can change predefined parts of the selected sentence) modes. The design of the speech generation user interface is similar to the design of our system for speech impaired people on mobile devices [3], but in VoxAid 2006 the data structures are realized in XML (eXtendable Markup Language). In the XML data files the categories and sentences of fixed and partly fixed text modes furthermore the editable parts of partly fixed text sentences are stored. Additionally - because VoxAid 2006 runs on desktop and laptop computers - the user input can be completed with the help of the mouse / touchpad and the keyboard more easily and more rapidly than on mobile devices.

The speech recognizer is handled on a separate part of the screen. The speech recognition window displays not only the recognized units, but the state of the call (e.g. dialing, in-call, etc.) and the interactions of the user on the other side of the line (e.g. DTMF signals) as well.

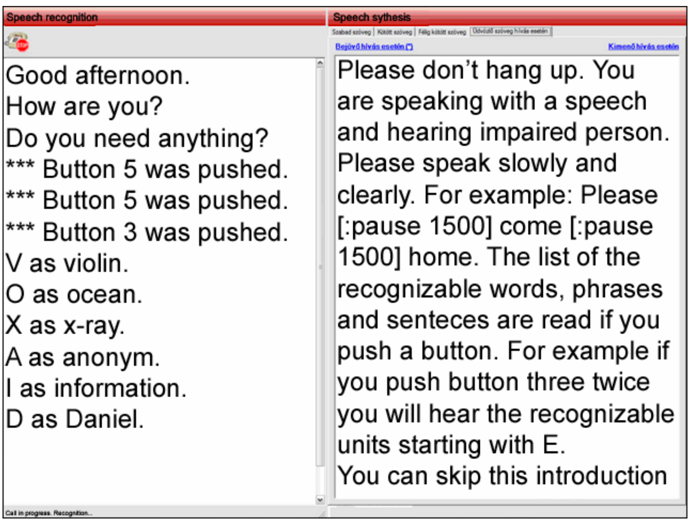


Fig. 2. The visual user interface. Speech recognition (left), speech synthesis (right, ‘welcome’ text in case of outgoing call).

The speech recognition engine is a speaker independent, isolated word and fixed vocabulary solution, as no continuous telephone speech recognition technology is available for the Hungarian language. Therefore a well designed vocabulary is required, and the possibility for the person on the other side of the line to find out which vocabulary elements can be recognized by the system. One can get help of the possible words, phrases and sentences with the buttons of the telephone. If one pushes a button on the other side of the phone line the words with the first letter, which is written on the button is read by VoxAid 2006. For example if s/he pushes button "4" three times, than the recognizable words, phrases and sentences starting with "I" will be read to the phone line. The user can stop the readout of the units by pressing any other button. With button zero the user can turn on/off the feedback feature. It is used to provide feedback of the results of the recognition process in order to help the speaking users on the other side of the phone line to learn the vocabulary. If the feature is turned on, VoxAid 2006 reads the recognized units or tells, if the recognition wasn't successful. After the user adapted to the speech recognition engine, for faster communication s/he can switch this feature off. There is also a possibility to tell unknown words (e.g. names of medicines) by spelling (e.g. V as violin).

In case of incoming and outgoing calls a 'welcome' text is read, which tells the user at the other end of the phone line that the called/caller party is not a computer, but a speech and/or hearing impaired person. The introduction can be skipped, if the user on the other side of the phone line pushes any button. The 'welcome' texts can be changed in VoxAid 2006.

The design and implementation of the ASR's vocabulary was quite a challenging task. From the users' point of view the ideal case would be if the vocabulary consisted of short words used daily, which can be learned by heart easily. We had to consider the recommendations of the ASR developers, however; to have better results the usage of units, which consist of more than 5-6 syllables, is recommended. During the design of the vocabulary we tried to keep the following principles:

1. Short units (1-4 syllables): We have chosen words and short sentences which have different spectral features. This way the speech recognizer has only a little chance to misrecognize the short units.
2. Longer units (5 or more syllables): We kept the same rule described in point 1. in case of longer sentences also, but we've faced the following problem: in case of sentences, where several choices are present the units are very similar (e.g. What do you want to eat for breakfast?; What do you want to eat for dinner?; etc.). To overcome the problem we rephrased the sentences in different ways (e.g. What do you want to eat for breakfast?; What do you prefer for dinner?; etc.).

4 Evaluation Results

The speech synthesis interface was originally designed for the first Voxaid [5] and used for more then ten years by a speech impaired person, and was further developed for SayIt [3]. With platform specific modifications the same interface is found in VoxAid 2006.

The most critical part of the system is the design of the speech recognizer's vocabulary. The probability of correct recognition is high, if the vocabulary consists of preferably 5-6 (or more) syllable long phrases, sentences. The disadvantage of longer units is that the people on the other side of the line cannot intuitively use VoxAid 2006's telephone interface, s/he always have to ask for help with the buttons or turn the automatic repetition on. Therefore a balance must be found between the length and the simplicity of the expressions.

The vocabulary was designed with the help of the speech recognizer developers and currently it contains 528 units. To test the accuracy of the speech recognizer an experiment was carried out. Twelve persons read out 50 randomly selected units in a noisy environment into an ISDN phone. At the other end of the phone line VoxAid 2006 was running. If a unit was not recognized or it was misrecognized the subject repeated it. If it still wasn't correctly recognized the subject had to repeat it again. After the third trial the subject had to read out the next unit, irrespectively of the result of the third attempt.

The probability of misrecognition and out of vocabulary event depends not only on the recognizer's accuracy, but on the design of the vocabulary as well. In consequence the test examines both the ASR accuracy and the suitability of the vocabulary. The results of the test can be seen in Figure 3.

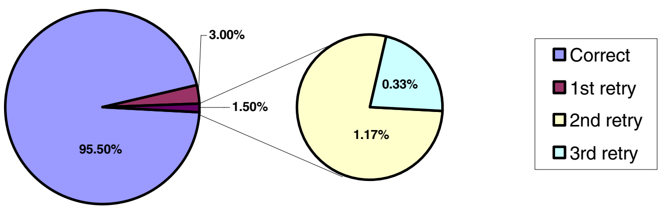


Fig. 3. The results of the test of the fixed vocabulary speech recognizer

The test showed that the probability of correct recognition of a unit out of 528 is 95.5%. Further 3% of the units were recognized during the second utterance (1st retry); 1.17% during the third utterance (2nd retry) and 0.33% during the forth utterance (3rd retry). We consider these results quite satisfactory; the ASR works well with the vocabulary we designed for the special task, although the speaking user at the other end of the phone line has to learn the recognizable units for rapid communication.

5 Future Plans

VoxAid 2006 is ready for testing. We are looking for speech and/or vocally impaired persons for cooperation to make user tests. After the user tests we shall ask the participants to fill out a form about the system and we also ask them to tell their preferences. According to these results we will tailor the visual and telephone user interface.

Nowadays mobile devices – primary the PDAs – have enough computing power for speech synthesis and recognition. There are many ways to port a speech synthesizer to a mobile phone with adequate performance [8], but there are problems with the performance of the speech recognizer. On the latest (with 520 Mhz CPU speed) PDAs the user independent speech recognizer performs well with about 200 units in the vocabulary. After the user tests are done in cooperation with the test subjects we shall try to scale the vocabulary for mobile devices.

6 Summary

VoxAid 2006 enables speech and vocally impaired people to communicate with voice over telephone. The communication can be 10-15 times slower than in case of non-disabled persons' conversation but the system enables speech and vocally impaired persons to perform everyday tasks with standard telephone users on the other end or to ask for help in case of emergency.

References

1. Telephone users with disabilities - the numbers, http://www.tiresias.org/phoneability/telephones/user_numbers.htm
2. Telecommunications for all - The Cost 219 project, Editor: Patrick R. W. Roe, ECSC-EC-EAEC, Brussels, Luxemburg, CD-90-95-712-EN-C, 1995
3. B. Tóth, G. Németh, G. Kiss, Mobile Devices Converted into a Speaking Communication Aid, Proc. of Computers Helping People with Special Needs, 9th ICCHP, 2004, Springer, 1016-1023
4. Hunnicutt, S. The development of text-to-speech technology for use in communication aids, in Applied Speech Technology, 1995, CRC Press, Inc., 547-563
5. G. Olaszi, G. Németh, VOXAID: An Interactive Speaking Communication Aid Software for the Speech Impaired, Proc. Eurospeech '93 Volume 3, 1993, 1821-1824
6. M. Karjalainen, P. Boda, P. Somervuo, T. Altosaar, Applications for the Hearing-Impaired: Evaluation of Finnish Phoneme Recognition Methods, Proc. Eurospeech '97 Volume 4, 1997, 1815-1818
7. M. Johansson, M. Blomberg, K. Elenius, L.-E. Hoffsten, A. Torberger, Phoneme recognition for the hearing impaired, Proc of Fonetik 2002, TMH-QPSR, Volume 44, 2002, 109-112
8. G. Németh, G. Kiss, Cs. Zainkó, G. Olaszy, B. Tóth, Speech Generation in Mobile Phones, Editors: D. Gardner-Bonneau, H. Blanchard, Human Factors and Interactive Voice Response Systems, 2nd ed., Springer, forthcoming

Evaluation of Effect of Delay on Sign Video Communication

Kaoru Nakazono[1], Yuji Nagashima[2], and Mina Terauchi[3]

[1] NTT Network Innovation Laboratories, Musashino-shi, Tokyo, 180 Japan
nakazono@core.ecl.net
[2] Kogakuin University
[3] Polytechnic University

Abstract. Evaluation tests of sign communication with delayed video are reported and the effect of delay on the communication is discussed. The authors constructed the delayed sign dialogue experimental system. Five kinds of tasks were assigned to deaf subjects and videos of performing the task with various delay times were recorded. By analyzing the data, sign communications was found to be more tolerant of the delay time than voice communication.

1 Introduction

The authors have been studying the evaluation of quality of sign language video supposing the achievement of sign telecommunication using videophones[1]. These studies were carried out with offline experiments in which encoded data was presented to subjects. These environments were far different from actual conversation that occurs over a videophone.

In real-time video communication, delay occurs easily at every procedure of picture encoding, transmission of data, and decoding of picture. Adverse impacts to communication are predicted due to this delay time.

The effect of delay on voice interaction has been studied in detail[2]. The effect of delay on voice interaction using videophones has also been studied[3].

Spoken Japanese was the language used in those studies. The effect of delay on sign language communication, which has not been studied so far, is predicted to be significantly different from that on auditory communication.

The authors constructed a delayed sign dialogue experimental system. Using this system, delayed sign communication can be recorded.

When the sign dialogue experiments are carried out, the method of communication can be regulated by assigning a task. In addition, smoothness of communication can be evaluated by measuring the completion time of the task. Determining the kind of tasks to be assigned is an important issue for this study. We assigned five tasks and carried out experiments with various delay times.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 659–666, 2006.

2 Method of Experiments

2.1 Delayed Sign Dialogue System

To simulate communication over actual videophones, a delayed sign dialogue experimental system was constructed, which can display the video of a companion on a prompter put in front of each subject. Subjects communicate by using sign language during the experimental procedure. The video image can be delayed using the digital video delay unit (SONY, VFD0401.) The layout of the system is shown in Fig. 1. Numbers or Japanese characters that indicate the content of tasks are superimposed on the prompter using the video titler. All conversations between subjects are recorded by digital betacam recorders.

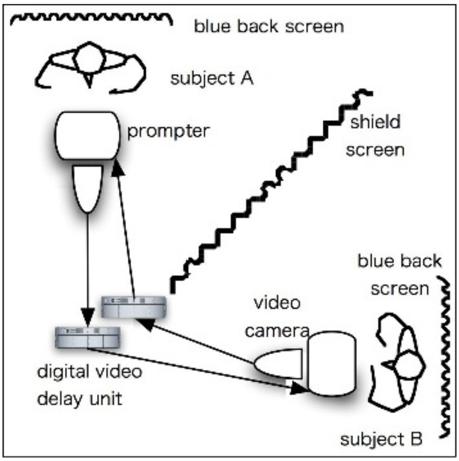


Fig. 1. Layout of Delayed Sign Dialogue Experimental System

Delay time can be controlled independently in either communication direction by multiples of a frame, from 0 (no delay) to 255 frames. In this study, delay times of both communication directions were set to the same value and controlled from 0 to 90 frames (3 seconds.)

Whereas the delay time on the communication network is generally shown by the round trip delay, we will unify the expression of the delay times on the basis of one-way trip delay. In the case of referring to data of other studies, we will divide the value of the delay by two.

2.2 Experimental Tasks

The following experimental tasks were carried out.

Alternating Count of Numbers. This task is counting from one to ten alternately between two subjects as fast as possible. If the subjects are signers, the numbers are shown by the number sign. If the subjects are not signers, the numbers are shown by the number of fingers of both hands.

Addition of One-digit Numbers. Tables of numbers are superimposed on the screens of both subjects. An example of these tables is shown in Fig. 2. A table with two columns and three rows is shown to the "subject A." Here, we call the left column "column A" and the right column "column S." A table with one column and three rows is shown to the "subject B". We call that "column B." Each number in "column S" is set to the sum of numbers of the same rows in "column A" and "column B," except for one number which is set to wrong number intentionally. All numbers in "columns A, B, and S" are set not to be greater than nine.

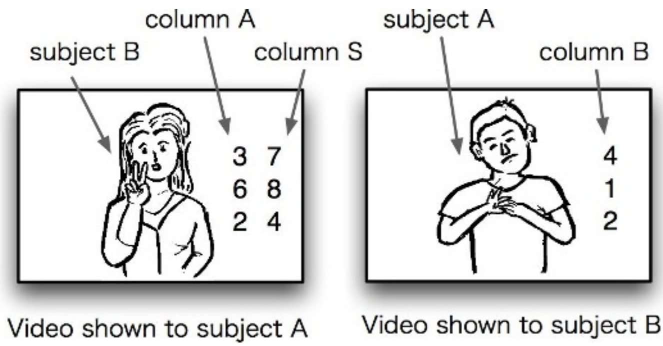


Fig. 2. Example of presented pictures of "addition of one-digit numbers" task

After the start signal, the subjects communicate and process the first numbers in the following way.

1) Both subjects synchronously express the first number in "column A" and "column B" respectively.
2) "Subject B" calculates the sum of the two numbers and expresses it. "Subject A" expresses the first number in "column S" that is synchronized to the sign expression of "subject B."
3) They verify whether their two numbers are identical.

Then, they process the second and third numbers in a similar way.

Checking of Numerical Table. Numerical tables (six digits × eight lines) are superimposed on the screens of both subjects. The tables are nearly identical but zero to two numbers of every line are randomly different. The subjects are instructed to check the different numbers.

Direction of Route. A map with many landmarks was given to a subject with the role of guide and an abbreviated map was given to a subject with the role of traveler. The guide is instructed to show the traveler the way to a goal.

Free Talking. The subjects can talk to each other without an assigned topic.

Table 1. List of categories and words

Category	Words
Numbers	1, 2, 3, 4, 5, 6, 7, 8, 9
Seasons	Spring, Summer, Autumn, Winter
Day of the Week	Monday, Tuesday, Wednesday, ...
Schools	Kindergarten, Elementary School, Junior High
Vehicles	Car, Bicycle, Train, Bus
Animals	Dog, Cat, Rat, Lion
Sports	Baseball, Football, Tennis, Swimming
Families	Father, Mother, Brother, Sister, Grandfather
Countries	Japan, USA, England, China
Fruits	Apple, Banana, Grape, Orange

Checking of Categories of Words. A list of words, which consists of ten categories, was prepared in advance. Each category contains four to ten words. List of words is shown in Table 1. The list is printed on a sheet of paper and shown to subjects prior to the experiments. During the experiments, two different tables of six lines of words are superimposed on the screens of both subjects. The subjects are instructed to check the categories of words line by line.

2.3 How to Analyze the Data

Effects of delay on performing tasks were evaluated in the following aspects: 1) observation of communication process, 2) temporal characteristics of signing, 3) completion time of the task, 4) lower limit of detection of delay.

Temporal characteristics of signing include duration of sign word expression, duration of transition between two sign words, frequency of nodding, and number of expressed words. These characteristics are analyzed manually by a signer.

The lower limit of detection of delay is measured by the double stimulus impairment scale method with a five-grade impairment scale. The reference condition with no delay and the test condition with various delays were presented alternately to the subjects. Every time after the trial under a test condition was finished, the subject was asked to evaluate the test condition keeping in mind the reference. The evaluation is based on the five-grade impairment scale with a description of each grade: 5-imperceptible, 4-perceptible, but not annoying, 3-slightly annoying, 2-annoying, and 1-very annoying. At the end of the series of sessions, the MOS (Mean Opinion Score) of each test condition is calculated by taking the average of the evaluated scores.

3 Results

3.1 Observation of Communication Process

The following tendencies were observed from the recorded video. The delay time that disturbs the communication differed by task. In the “checking of numerical table" task and the "direction of route" task, process of communication was frequently disturbed when the delay time exceeded 500 milliseconds. On the other hand, in the "free talking" task, some pairs of subjects were not aware of delays longer than one second.

For the task in which subjects needed to consult information printed on papers like maps in the "direction of route" task, a problem occurred that is intricately related to delay time.

To consult the map, the subject needed to drop his eyes from his companion on the screen. Unlike auditory communication, signer misses to understand what the companion signing if he looks away. In the case of no delay, the companion can immediately notice that the subject is looking away. The companion can wait until the subject looks again or waves hands to get his attention back.

However, when the delay time becomes longer than some interval, the subject frequently fails to notice that the companion is looking away. In such a manner, sign communication that requires consulting printed information can be disturbed.

This kind of task is not suitable for quantitative analysis of the effect of delay.

3.2 Temporal Characteristics of Signing

Temporal characteristics of signing were analyzed in the "checking of numerical table" task. The duration of sign expression became longer for some sign words with the lengthening of the delay time. This long duration is caused by sustaining the hand sign while waiting for the delayed action of the companion. Excluding these sustained hand signs, distributions of sign expression durations were almost identical regardless of delay.

Then, average numbers of expressed words per minute were determined by a signer by looking at a recorded video of the "free talking" task. The result is shown in Fig. 3. The duration of each trial was about three minutes. The result indicates that the average number of expressed words was not affected by delay time.

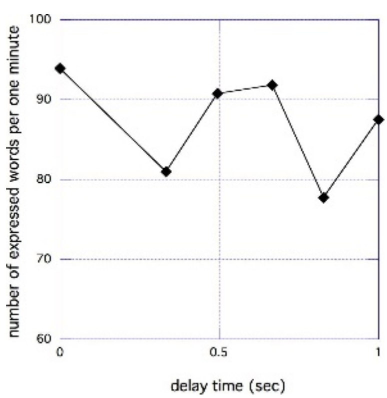


Fig. 3. Number of expressed words per minute for "free talking" task

3.3 Completion Time for the Task

Positive correlations were observed between delay time and completion time for the "checking of categories of words" task, "alternating count of numbers" task, and "addition of one-digit numbers" task. The result of two pairs of subjects performing the "checking of categories of words" task is shown in Fig. 4

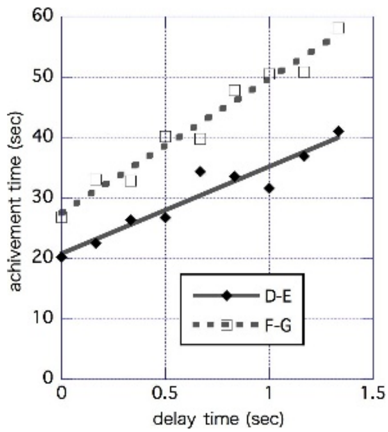


Fig. 4. Completion time and delay for the "checking of categories of words" task

In the delayed sign communication, a picture of the companion is presented with a constant delay time. For the achievement of the task like "checking of categories of words," subjects cannot go ahead until they receive information from their companions. Accordingly, the fact that the relationship between delay time and completion time is linear indicates that the frequency of information exchange is fixed. The fact that the gradients of lines are different depending on the pair of subjects indicates that the frequencies of information exchange, namely the procedures of problem solving, are different. Furthermore, the fact that the linearity is maintained even if the delay time is extended to 1.33 seconds indicates that the communication is not disrupted by the delay.

3.4 Lower Limit of Detection of Delay

MOSs of the "alternating count of numbers" task and the "addition of one-digit numbers" task for various delay times are shown in Fig. 5. All of the data degrade with the lengthening of the delay time. In the "alternating count of numbers" task, MOSs of two methods of communicating numbers were almost the same. Degradation of the MOS of the "addition of one-digit numbers" task is more gradual than that of the "alternating count of numbers" task. That finding implies that the load of calculation blocked the awareness of delay.

Then, the lower limit of detection of delay was investigated. The graph of the percentage of maximum scores of the two tasks for various delay times is shown

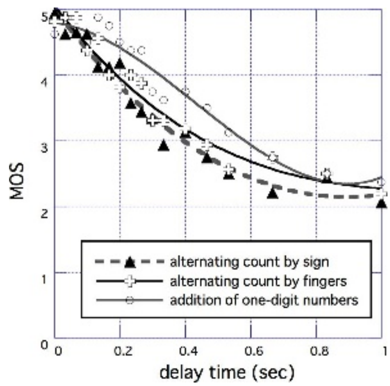


Fig. 5. MOS and delay time

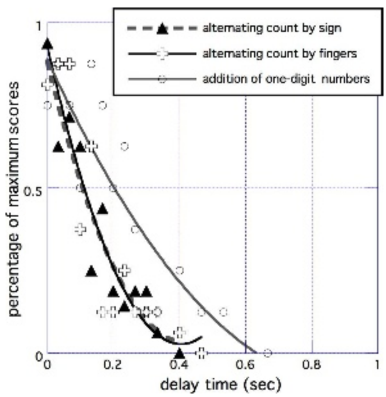


Fig. 6. Percentage of maximum points and delay times

in Fig. 6. The lower limit of detection of delay is defined as the time in which the percentage of maximum scores degrade less than 50 percent.

From the graph, the lower limit of the detection of delay for the very simple "alternating count of numbers" task was detected to be 100 milliseconds regardless of the numerical method. The limit for the "addition of one-digit numbers" task was 200 milliseconds.

The limit for detection of delay in voice communication has been reported to be 45 milliseconds [2]. These values for visual communication are relatively long compared to those of voice communication.

4 Discussion and Conclusion

The aim of this study was to determine the effect of delay on sign communication, which is a critical issue in sign communication using videophones. To achieve this aim, five communication tasks using sign language with various delay times were carried out.

From the analysis of these experimental data, sign communication proved to be more tolerant of the delay time than voice communication.

The following three hypotheses can be proposed as reasons for the tolerance to delay:

1. There are differences in physiological and cognitive mechanisms between auditory and visual perception.
2. Sign communication is tolerant of simultaneous signing whereas simultaneous speech is avoided in voice communication.
3. End of turn can be easily predicted in sign communication. Therefore, turn taking can be established smoothly.

From the result in section 3.4, the lower limit of detection of delay in sign communication was found to be much longer than that of voice communication in the very simple "alternating count of numbers" task. This finding is still valid for trials of hearing people showing number by fingers of both hands. This result supports hypothesis 1).

Furthermore, from the result of the latter part of section 3.2, deaf people seem to start signing without waiting for the end of their companion's utterance. This observation supports hypotheses 2) and 3).

In this paper, we focused on the temporal characteristics of manual sign. It will be of much interest to investigate the emergence of non manual sign such as the facial expression or nodding. Furthermore, the reason for the tolerance of delay in sign communication will be studied by refined experiment tasks.

Acknowledgement

We thank the native signers who supported us in recording and analysis.

This report is supported by Japanese Grants-in-Aid for Scientific Research (A)(1) #14208030 and Grant-in-Aid for Scientific Research on Priority Areas #16091209.

References

1. K. Nakazono, Y. Nagashima and A. Ichikawa: "Digital encoding applied to sign language video (to be issued)", IEICE Transactions on Information and Systems, (2006).
2. N. Kitawaki and K. Itoh: "Pure delay effects on speech quality in telecommunications", IEEE Journal on Selected Areas in Communications, **9**, 4 (1991).
3. T. Kurita, S. Iai and N. Kitawaki: "Effect of transmission delay in audiovisual communication (in Japanese)", The IEICE Transactions BI, **J76-B-I**, 4, pp. 331–339 (1993).

Japanese JSL Translation and Searching Display Conditions for Expressing Easy-to-Understand Sign Animation

Sumihiro Kawano[1], Chie Izumi[2], Takao Kurokawa[2], and Kazunari Morimoto[2]

[1] Faculty of Industrial Technology, Tsukuba University of Technology, Amakubo, Tsukuba 305-0005, Japan
kawano@a.tsukuba-tech.ac.jp
[2] Graduate School of Science and Technology, Kyoto Institute of Technology, Matsugasaki, Sakyo-ku, Kyoto 606-8585, Japan
{chiexx05, kurotak, morix}@kit.jp

Abstract. This paper described a bi-directional translating system between Japanese and Japanese Sign Language and two experiments conducted to clarify necessary conditions for displaying easy-to-read animation of a person model speaking sign language synthesized by the system. In Experiment 1 hearing-impaired and hearing subjects judged identity of a pair of animation. Resolution and frame rate had an main effect against correct answer rate. While the hearing had tendency to watch the whole body of the model, the hearing-impaired focused their attention chiefly on handshape and movement. In Experiment 2 the hearing-impaired read signs on animation. For correct reading above 90 % sign animation needed to be displayed with more than 10 x 8 cm, 113 x 90 pixels and 8 f/sec. This condition was much relaxed compared with that in Experiment 1.

1 Introduction

Computer-synthesized sign animation is one of promising means for assisting communication of the disabled. It can be used in various purposes including translation between different sign languages, translation between sign and verbal languages, and news programs. Many systems have been proposed and are under development for sign translation and animation synthesis. The authors are also developing a system for bi-directional translation between Japanese and Japanese Sign Language (JSL) [1, 2]. Most of such systems including ours employ animation featuring a person model that speaks sign language under control of a computer.

The hearing-impaired, however, are not accustomed to read sign language presented in animation. Some people felt something different about it and other people felt difficult to read the model-speaking sign language. Besides, a very small number of people with hearing difficulty believed that animation could not express sign language like native signers. Based on such feedback from signers and sign language

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 667 – 674, 2006.

interpreters the authors and collaborators have repeatedly made many substantial improvements on our person model and the translating process. However, there are some remaining problems about structure and functions required for person models and physical conditions for displaying sign animation. As to the latter case we have to take a wide variety of screens into consideration.

This paper briefly introduces our Japanese JSL translating system under development and describes the results of two experiments conducted to search necessary conditions for displaying easy-to-understand JSL on a screen.

2 Japanese JSL Translating System

Our system is outlined in Fig. 1 [1]. It translates JSL into Japanese and Japanese into JSL, and has the aim to assist mutual communication between a hearing-impaired person and a hearing person who does not understand JSL. This system is oriented to an intermediate type of JSL that is in widespread use through TV media in Japan. The system employs a case frame as an intermediate expression between Japanese and JSL in both directions of translation. For translation of JSL into Japanese we use finite automata [3]. Now this remains in the stage of JSL word recognition.

Another function of the system is to translate Japanese into JSL as shown in Fig. 1. It proceeds as follows. An input Japanese sentence is processed by a morpheme analyzer Juman for splitting it into morphemes and a parser KNP (both programs were developed at Language Media Laboratory of Kyoto University), and is represented by a case frame which describes its semantic structure with case relations between verbs and noun phrases, tense, aspects and modalities. Then the case frame is transformed into a sign expression, or a sequence of sign morphemes, via a Japanese-JSL dictionary. Using the files called "talk files" which contain CG notations, or sets of codes assigning skeletal segment configuration, facial expression and mouth shape for animating the person model, the sign expression is visualized by means of the synthesis-by-rule method using grammatical rules established by us. The outcomes of translation, or JSL sentences, are presented on a screen with animation featuring the person model.

Our person model has a female figure with a simplified skeletal structure and wears multi-polygonal skin and blue clothes, standing against a green background as shown in Fig. 2. Its skeleton has 42 virtual joints and its facial expression and mouth shape are controlled by moving a small set of polygon vertices. The model has been repeatedly improved to express natural and easy-to-read sign language [2, 4, 5]. Though grammar of the intermediate type of JSL has not been well known, we analyzed a lot of its sentences and syntactic and semantic rules were derived [6, 7]. Many of them have already been installed into the system. They include rules of nonmanual actions such as eyes, eyebrows and facial expressions. Some of them are semantic and are described in the talk files in advance. Because others carry syntactic information on sign phrases and sentences, they are applied to movements of the model in the last stage of animation synthesis [8, 9]. Through experiments in which the hearing-impaired participated the correct recognition rate of JSL sentences with facial expressions and nodding was found to be 10 % higher than those without facial expressions and noddings.

Fig. 1. The system bi-directionally translating between JSL and Japanese. The shadowed part illustrates translation from Japanese into JSL.

Fig. 2. The person model and examples of sign animation

3 Experiment 1: Image Matching

Translation of Japanese into JSL is desired to be in practical use in many scenes such as counters in hotels and banks, telephone-mediated communication, interaction with a doctor at hospitals and news broadcasting. In actual use the system should take advantage of a variety of screens and communication facilities. As to screens we have to consider a variety of them from a very small one on a mobile phone to a very large wall surface screen. We, however, do not know a range of physical parameters of display size, resolution and frame rate necessary to display easy-to-understand JSL animation.

As far as video images of signs, B. F. Johnson and J. K. Caird showed that 1 or 5 f(frames)/sec was amply sufficient for beginners of signing to read learned ASL signs [10]. Kamata et al. [11] found that understanding rate and subjective preference significantly decreased at screen size smaller than 2 inches. We, however, have no information on how high quality is needed on display in the light of human perception, recognition or cognition of sign animation as mentioned above.

Another purpose of this paper is to show the solution of this problem. For this we carried out two experiments; Experiment 1: image matching and Experiment 2: sign understanding. Experiment 1 was planned to know display conditions necessary for people to judge identity of two pieces of animation. In Experiment 2 hearing impaired persons read JSL signs from animation displayed under various conditions. Because signs had clues that enabled to understand their meanings, necessary conditions to understand signs were expected to be looser than in Experiment 1.

3.1 Method

In Experiment 1 a series of image matching trials were imposed to subjects. In a trial two pieces of animation (duration: 0.9-2.5 sec) were presented at a short interval and subjects were asked whether those were the same or not. After they judged that the two were not identical, they were required to answer what was/were different. The foregoing animation was taken from five JSL signs and was displayed with 15 (longitudinal) x 12 (transversal) cm in size (A in Fig. 3), 700 x 500 pixels and 15 f/sec. In contrast the following one was taken out of about 300 pieces of animation. It was either the same as the foregoing one or differed from it in one or two items among the hand path, hand shape, eye opening and mouth shape change, and was presented under one of 27 display conditions combining three display sizes (a, b and c in Fig. 3 in Table 1), three sorts of resolution and three frame rates listed in Table 1. The subjects could watch the second animation up to three times if they wanted.

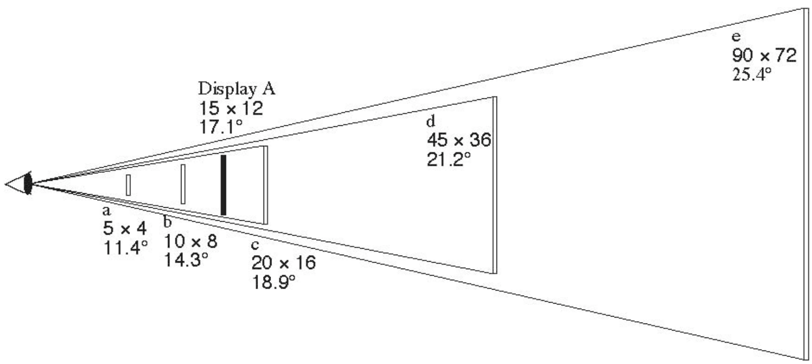


Fig. 3. Display size and vertical visual angle used in Experiment 1(A, a-c) and Experiment 2 (a-e). Display size and visual distance are drawn in the same scale.

Table 1. Display conditions for Experiments 1 and 2

display name	a	b	c	d	e
Display size (cm)	5×4	10×8	20×16	45×36	90×72
visual distance (cm)	25	40	60	120	200
vertical vision angle (°)	11.4	14.3	18.9	21.2	25.4
Resolution (pixels)	113×90	225×180	450×360		
Frame rate (f/sec)	4	8	15		

10 hearing-impaired students (5 males and 5 females) and 18 hearing students (12 males and 6 females) participated in the experiment. All of them had no experience of watching sign animation before. One session of the experiment consisted of 10 to 25 trials under the same display condition. The trial order was randomized among the subjects.

3.2 Results

We focused on the items having content different between the paired pieces of animation, because we thought that the subjects answered "same" when they were not assure of difference between the two. Fig. 4 shows the ANOVA, whose factors were size, resolution and frame rate, and it revealed that there were no interaction among the three factors and between any two factors. Significant main effects were found in resolution ($F = 7.981$, $p < 0.01$) and frame rate ($F = 3.719$, $p < 0.05$) of the hearing-impaired group and in resolution interaction ($F = 19.882$, $p < 0.001$) of the hearing group. Significant differences between the two groups are indicated in Fig. 4. These results tell that display size had no effect against correct answer rate. However, the main effect was due to resolution and frame rate.

The two groups showed different correct answer rates in each of items having difference in pairs of animation as shown in Fig. 5. From this we may say that the hearing-impaired and the hearing had different characteristics in attention. The hearing subjects had tendency to watch the whole body of the model. In contrast, the hearing-impaired focused chiefly on hand shape and movement, reflecting their attention given to particular regions that convey linguistic information.

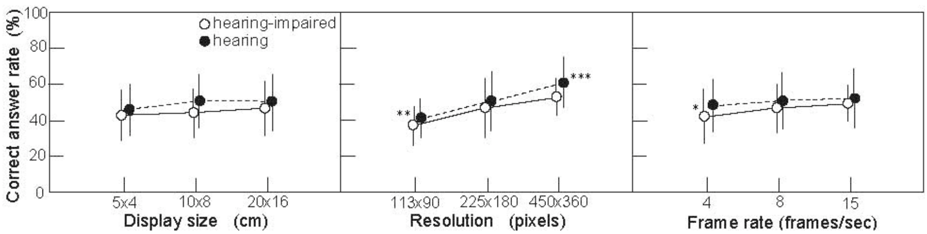


Fig. 4. Correct answer rate as a function of display size, resolution and frame rate. Significant level: *** p < 0.001, ** p < 0.01, * p < 0.05.

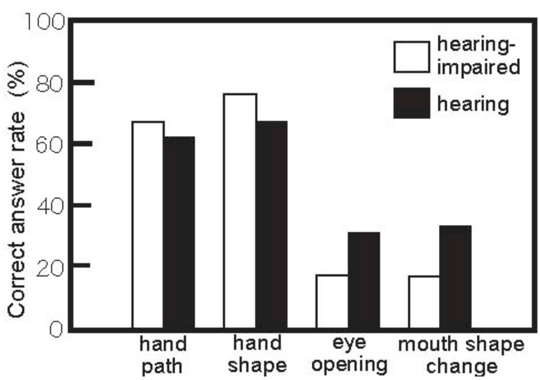


Fig. 5. Correct answer rate in the hearing-impaired and the hearing groups with respect to items that have difference in pairs of animation

4 Experiment 2: Sign Understanding

4.1 Method

In this experiment JSL signs were presented to subjects and they were asked to write their meaning. 13 signs listed in Fig. 6 were selected so that they had a main key or clue to recognize and understand in any one of the items; (1) mouth shape and relation between a hand and the mouth, (2) handshape and/or a shorter hand trajectory, (3) movement of both hands with a relatively long trajectory. For example "red" and "white" have the same hand movement, but the mouth is largely open in "red" and open a little in "white". However the teeth are shown only in "white". "Sweet" and "sugar" differ in the number of moras, therefore change of mouth shape serves as an important key. When a main key is highly visible, people can easily catch its meaning and will respond correctly even under wrong display conditions. But when they have to watch a small region or subtle movement of the model body, a contrary phenomenon will occur. Each sign animation was successively presented twice on one of the 45 display conditions (five display sizes a-e, three sorts of resolution and three frame rates) in Fig. 3 and Table 1.

Subjects were 16 males and 21 females between the ages of about 20 and 60. They all had over 10 years of experience in using JSL but had not watched sign animation. Each subject had 25 to 35 sessions. One session consisted 4 to 8 trials under the same display condition. The order of the trials was randomized among the subjects.

4.2 Results

Fig. 6 shows correct answer rate for each of 13 signs. These signs can be classified into three groups according to the item that contains a main key for understanding the signs as mentioned in **4.1**; say (1) mouth signs, (2) one-handed signs, and (3) two-handed signs. Mean correct answer rates were smallest in the mouth signs and largest in the two-handed signs although all were above 80 %. This result reflects easiness for perceiving the key as described in **4.1**. Fig. 6 integrated the data in all 45 display conditions and shows that sign reading was very much easier than image matching.

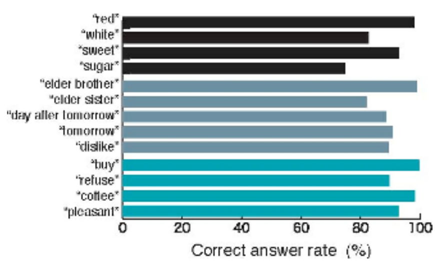


Fig. 6. Correct answer rate for each sign used in Experiment 2. The upper four signs are mouth signs, the middle five are one-handed signs and 117the bottom four two-handed signs.

The above result suggests that signs with a subtle key are difficult to read even under better conditions in display size, resolution and frame rate. In order to confirm that this is true, we calculated the values of the three display parameters necessary to attain correct answer rates of 80, 90 and 100 % in the three sign classes. The correct reading rates in percent are illustrated in the 3D parameter space (size × resolution × frame rate) based on the calculated values as in Fig. 7. The space above the surface marked "*r* %" means a set of parameters with which signs more than *r* % was correctly understood.

From this figure as well as Fig. 6 we can know that the mouth signs needed the most severe conditions and that the two-handed signs were easiest to read. For example the mouth signs could not be read with correct answer rate 90 % under 5 x 4 cm or 113 x 90 pixels. Even in the two-handed signs correct answer rate was lower than 90 % when resolution was 113 x 90 pixels and 4 f/sec.

These results showed that (i) reading signs with a larger trajectory was not influenced by display size, (ii) signs having hand shape and/or a smaller trajectory as keys were intricately influenced by the three parameters, and (iii) mouth signs were strongly affected by the three parameters. The most rigid conditions were required in reading mouth signs.

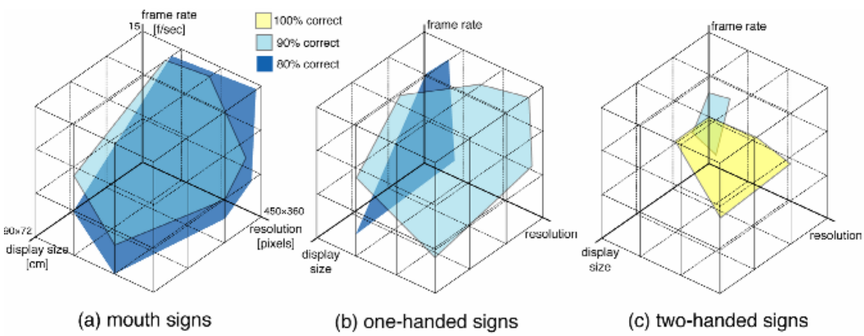


Fig. 7. Necessary display conditions for 80, 90 and 100 % correct reading of the three kinds of signs

5 Discussion

Our system worked well in synthesizing sign animation. All the subjects favorably accepted the animation, although they experienced it for the first time. However, the authors strongly think it to be improved for synthesizing sharp sign images.

As expected the conditions necessary for image matching and reading signs were entirely different: sings were correctly read under looser conditions than in image matching. The two experiments were different in cognitive task and load which subjects needed. In Experiment 1 subjects were required to memorize whole of animation for checking at the following step. For the hearing-impaired the animation was not relevant to sign language, and this point was the prime reason why the results were different between the two experiments.

Display conditions necessary for reading signs were varied according to easiness of perceiving keys. The display size of 4 x 5 cm that assumed mobile phone screens was unsuitable for displaying single signs, especially signs with subtle keys. But in case of presenting sign sentences small screens are of great promise for mobile devices. If we expect correct reading above 90 %, sign animation should be displayed on condition better than 10 x 8 cm, 113 x 90 pixels and 8 f/sec. We think that these values may have applicability to other sign languages.

6 Conclusions

Using the Japanese JSL translation system under development we carried out two kinds of experiments. The results revealed that physical matching of animation needed image quality higher than sign reading. Though necessary condition was obtained for displaying signs in general, they were preliminary toward the final goal of determining ones for presenting sign animation in practical scenes. Optimal conditions should be determined by taking the case of sign sentence reading into consideration.

Therefore, the next natural step is to carry out reading experiments using a variety of JSL sentences. We can predict that the necessary condition given in this paper will be relaxed in reading sentences because people can use more keys and context information embedded in sentences.

Acknowledgement. The authors thank Prof. K. Kanda of Chukyo University for his proposal of signs for Experiment 2 and his suggestion in preparing this manuscript. A part of this study was supported by the Scientific Research Fund (No. 16091203), Ministry of Education, Culture, Sports, Science and Technology.

References

1. Kurokawa, T., Morichi, T.: Watanabe, S., Bidirectional Translation between Sign Language and Japanese for Communication with Deaf-Mute People. Advances in Human Factors/Ergonomics, Vol. 19B. (1993) 1109-1114
2. Kawano, S., Kurokawa, T.: Development of a System Translating Japanese into Japanese Sign Language. J. Japan Soc. Fuzzy Theory and Intel. Informatics. 16 (2004) 485-491
3. Fujishige, E., Jan, K., Kurokawa, T.: Recognition of Japanese Sign Language Words Using Finite Automata. Human-Computer Interaction, Vol. 2. (1999) 167-171
4. Miyashita, J., Iwata, K., Morimoto, K., Kurokawa, T.: Improvements in Motor Control Methods of a Human Model for Sign Animation and Their Effects. Proc. Human Interface 2003 (2003) 305-308
5. Kurokawa, T.: Representation of Sign Animation for Japanese-into-Japanese Sign Language Translation. J. Visualization Soc. Japan. 24, Suppl. No. 1 (2004) 273-276
6. Kawano, S., Takahashi, Y., Kurokawa, T.: An Analysis of Blinking, Gazing, Nodding and Their Meanings in Sign Language Interpretation of TV News. Human Interface News and Report. 11 (1996) 435-440
7. Demura, T., Itou, T., Fujishige, E., Kurokawa, T.: Analysis of Deictics in Japanese Sign Language by Using a Database of JSL Sentences. Rep. 3rd IEICE Technical Meeting for Sign Language Informatics (2000) 37-40
8. Kawano, S., Kurokawa, T.: The Effects of Facial Expression on Understanding Japanese Sign Language Animation. Human-Computer Interaction, Vol. 1. (1999) 783-787
9. Kawano, S. and Kurokawa, T.: Facial and Head Movements of a Sign Interpreter and Their Application to Japanese Sign Animation. Proc. ICCHP 2004 (2004) 1172-1177
10. Johnson, B. F., Caird, J. K.: The Effect of Frame Rate and Video Information Redundancy on the Perceptual Learning of American Sign Language Gestures. Proc. CHI '96 (1996) 121-122
11. Shionome, T., Kamata, K., Yamamoto, H.: A Consideration on Relationship between Sign Language Perception and Screen Size. Tech. Rept. of IEICE, HCS2003-61 (2003) 29-34

Design and Development of Several Mobile Communication Systems for People with Hearing Disabilities

Jose L. Martín[1], Sira E. Palazuelos[1], Jerónimo Arenas[2], Javier Macías[3], and Santiago Aguilera[4]

[1] Department of Electronics, Polytechnic School, University of Alcalá, Campus universitario s/n, 28805 Alcalá de Henares, Madrid, Spain
{jlmartin, sira}@depeca.alcala.es
http://www.depeca.uah.es

[2] Department of Communication Technologies, Polytechnic School, Carlos III University
jarenas@ieee.org

[3] Speech Technology Group, Department of Electronics Engineering, E.T.S.I. de Telecomunicación, Technical University of Madrid
macias@die.upm.es

[4] Department of Photonics, E.T.S.I. de Telecomunicación, Technical University of Madrid Madrid, Spain
aguilera@tfo.upm.es

Abstract. During the last decade we have attended to an impressive development of mobile communications which, unfortunately, deaf and hearing-impaired community cannot (in principle) take advantage of. In order to favour this people to take part in the Information Society, we have designed and developed some PC-based systems and applications which will provide several text-based services to them, such as real time and text mode communication between mobile text telephones and those connected to the PSTN or direct accessibility to Urgency Call Centers. We emphasize in this article the advantages of the software design methodology followed, which has led to the implementation of two systems which have shown to be robust and versatile in operation.

1 Introduction

The invention of the telephone in 1876 by Graham Bell, improved the social welfare of millions of people, by allowing the communication between distant places. At first, this communication system could not be used by the hearing-impaired people; however, the development of the first text telephone terminals in the seventies changed this scenario. In these devices the microphone and receiver of a conventional telephone are substituted for both a display and a keyboard, allowing real time and text mode conversations. Although there are different models of text telephones (we will refer in the rest of this article to the DTS, the most widespread terminal in Spain), all of them present a common drawback. As

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 675–682, 2006.

they don't use standard communication protocols, only communications between terminals of the same kind are feasible.

Moreover, with the appearance and increasing significance of mobile communications one more challenge joins in the final objective of a global communication system completely integrated for the deaf community.

Bearing these ideas in mind, a standard mobile telephone (Nokia Communicator) was provided with a software application in order to allow deaf users to maintain a text mode conversation between two of these terminals, or one of these and a PC connected to the standard PSTN network via modem.

However, there were still many facilities which could be implemented to further improve the telecommunication capabilities of deaf people. In this direction, we have face two new aims that will be described in this paper:

– making the communication between a mobile terminal and the conventional DTS possible
– and providing Urgency Call Centers with an integrated communication system to receive calls from Nokia Communicators.

In order to implement these functionalities, we decided to develop two different and separated solutions: an Automatic Interworking System (AIS) and a dynamic library (SUMMA061). **AIS** is a complex PC based system with the functionality needed to establish real time and text mode communication between Nokia Communicators and DTS terminals from end to end (see Fig. 1). **SUMMA061** is an specific dynamic library designed to be integrated in the emergency Call Center standard software to attend Nokia communicator calls (see Fig. 2).

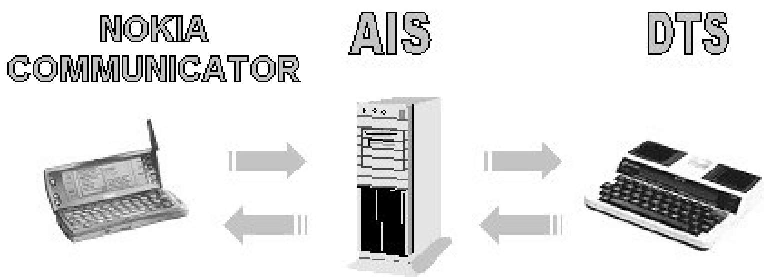


Fig. 1. Communication between Nokia Communicator and DTS, through the AIS

Though the Nokia Communicator-DTS communication was at first our main concern in the conception of what AIS should be, we soon realized that this system could host many more communication services for deaf people. By using an structured analysis and design methodology ([1]) (summarized in Table 1) we have tried to develop both systems in which new functionalities can be added in an easy and modular way.

Fig. 2. Communication between Nokia Communicator and emergency Call Center, through the AIS

Table 1. Structured Analysis and Design Methodology

Step	Description
1.	User Requirements Analysis
2.	Hardware Design
3.	Software Design (leading to a modular software)
4.	Implementation of the System
5.	Proof of the different modules
6.	Proof of the whole system by the users

2 System Requirements

Requirements imposed by users for the communication between the DTS and the Nokia communicator were:

- AIS must provide text conversations between DTS and Nokia Communicator without modifying (unless strictly necessary) the actual user interfaces (so the user doesn't need to get used to new protocols for making calls). Any modifications made, must be very easy to understand and manage.
- Telephone calls might be originated and released by any of the sides of the conversation.
- AIS must provide resources for up to 30 simultaneous conversations in real time.

For the addition of new communication services to AIS:

- AIS must have a modular software design which facilitates the incorporation of new services offered to the hearing-impaired community.
- The addition of new facilities must be carried out in a systematic way.

And finally, requirements for the SUMMA061 library were:

- The application must be absolutely robust to attend any kind of emergency call without hanging up.
- SUMMA061 should be fully integrated in the standard call center software (that keeps medical data, personal treatment information, etc. depending on the kind of emergencies attended at the call center).

- The application should allow encoded recording of every call, storing the number phone, length and content of the conversation for later analysis.
- Application access should be protected by several access levels permissions.
- Lost calls must be easily recover.

3 AIS Development

3.1 Hardware Design

AIS hardware architecture must supply, at least, means of communication with terminals both connected to the GSM and PSTN/ISDN networks. However, as we will discuss, a more general solution has been adopted. Mobile telephones will access AIS through TCP/IP sockets, which provide them with a connection oriented path, guaranteeing the real time requirements at this side of the conversation. This was not the only available solution, but it has the advantages of cheapness (just a LAN board with an IP address are needed), capacity (there is not a limit in the number of mobile terminals which can be connected to the AIS at a time), simplicity (lots of programming tools with sockets support are available, even at zero cost) and versatility (the Internet connection can also be used for other purposes).

For communication with DTS, AIS will use an ISDN primary access (E1), which allows up to 30 simultaneous calls. This interface will be managed by a Dialogic D/300-PCI board. Dialogic API implements useful high level routines ([2]) for dialing a number, putting the line on hook / off hook, playing a file, or recording the sound received, simplifying enormously the programming tasks.

Fig. 3 shows a non exhaustive communication scenario resulting from the decisions discussed above. It represents not only the terminals involved in a DTS-Nokia 9110 conversation, but also other terminals that can get to the AIS through the telecommunication networks depicted (for future communication services).

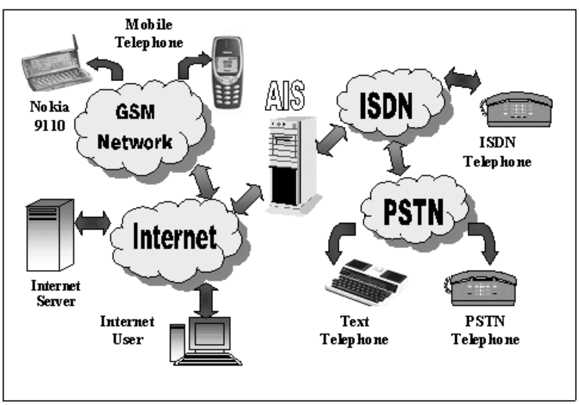


Fig. 3. Communication possibilities of AIS

3.2 Software Design

AIS has been designed using a descendent splitting scheme: in a first step we split the functionality we want the AIS to implement, into well differentiated units. Next, we divide each unit into smaller ones until we get modules small enough to easily program them. The first splitting results in the three blocks:

The Control Unit is in charge of managing all the resources of AIS (memory, B channels on ISDN interface, TCP ports, etc). In principle, the AIS is to attend up to 30 end to end communications of any type (DTS-Nokia Communicator must be treated just as a possibility at this point of design), each of then requiring (at most) a B channel, a TCP port, and some memory space for its exclusive use. As a consequence of this, we have further divided the Control Unit into three modules:

– Channel Information Table (TCYP), which will keep information about the 30 communication channels: their availability (free / in use), telephone numbers of the users, and any other relevant information.
– The Buffer Space (TB), is a table with 30 blocks of memory, one for each communication channel.
– Tables Manager Module (GT). It must maintain the coherence in the system at any time, by dynamically assigning or releasing the resources. It is the only module which can access the TCYP, and therefore, the most critical module in the AIS.

The Mobile Interface Manager Unit is made up of three kinds of modules:

– Communication Request Manager (S255). This module is permanently waiting for a new user to arrive at the Internet side of AIS. When this happens, it informs the GT about the event, and the service the user has demanded.
– Service Dependent TCP/IP Thread (HT). These threads are the ones that implement the functionalities required by the communication service demanded by the user. There will always exist one of them for each communication taking place at the system.
– TCP/IP Threads Manager (GHT) will dynamically create and destroy the HT threads.

The ISDN Interface Manager Unit is similar. The names we have given to its modules are listed next:

– Incoming Calls Manager (GLLE).
– Software Dependent ISDN Threads (HR).
– ISDN Threads Manager (GHR).

Fig. 4 shows the full splitting of the software, and the relations between the modules in it.

We have chosen C++ to implement the system. The reason is that we needed a language with the following characteristics ([3]):

– multithread programming: in AIS several tasks must run simultaneously, and each module in the architecture will be implemented with its own thread.

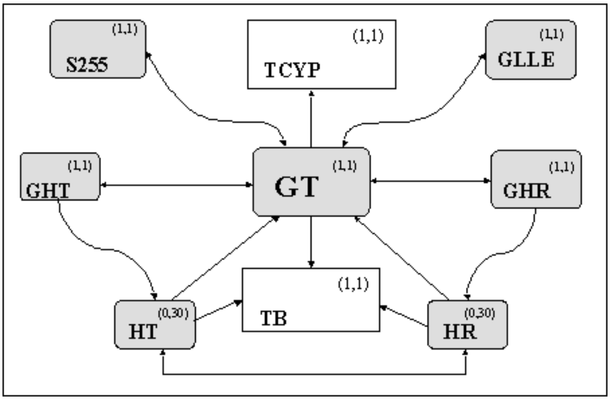


Fig. 4. Full AIS software architecture

- object oriented programming: so each module can be implemented as a different object with its own variables and methods.
- messages communication: the modules interchange messages in order to cooperate asynchronously.

3.3 Implementing DTS-Nokia Communicator Communication

As an example of the utility of AIS, we briefly discuss how it has been used to make it possible the communication between DTS and Nokia Communicator terminals. In order to implement it, we have developed two instances of HT and HR, called HTO and HRO, respectively.

An HTO thread simply receives characters from the socket that keeps it in communication with a mobile telephone. Next, the HTO writes this characters in the space of the TB assigned to the conversation. At the same time the HTO is checking the TB for the availability of characters received from the DTS by the HRO.

The HROs are more complex in the sense that they must implement routines to demodulate the characters received from the DTS terminals, and modulate the characters the Nokia Communicator has sent to the DTS.

Finally we want to remark that we have already added a new service to AIS which allows the sending of synthesized messages written by deaf people (using Nokia Communicator) to non deaf people (using standard telephones). The messages are synthesized in AIS by HRC, one of the two new kinds of threads which had to be designed and implemented in order to create this service (also operative at present).

4 SUMMA061 Development

4.1 Hardware Design

SUMMA061 is installed in the Call Center PC's and only needs a standard modem to be connected to the ISDN network.

4.2 Software Design

Software design of SUMMA061 is simpler than AIS since this dll will only manage one call. Even so, different concurrent threads will be necessary to handle the user interface and modem issues simultaneously.

Fig. 5 shows principle threads involved in the application:

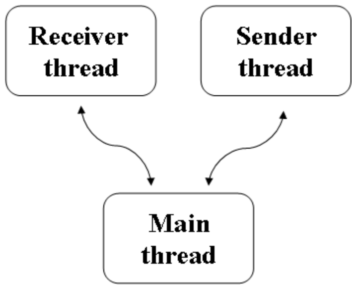


Fig. 5. Simplified SUMMA061 software architecture

- Main thread is in charge of coordinating user actions, user interface appearance (see Fig. 6) and information received or sent to receiver and sender threads.
- Receiver and sender threads communicate with the modem and controls receive and send buffers.

These last two threads are kept asleep, saving memory and CPU resources, until any event is produced in the modem or by the main thread. Communication between those treads and the main one is implemented by means of messages as it is in the AIS model.

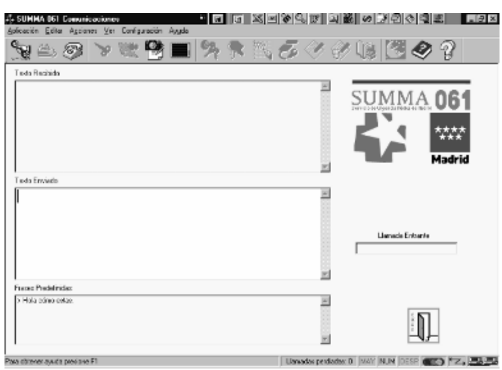


Fig. 6. SUMMA061 appearance

5 Conclusions and Future Work

The systems we have presented have already been tested and have shown to be very robust in operation. We have used AIS for providing DTS- Nokia Communicator communication, and the possibility that deaf people send synthesized messages to non deaf ones. The addition of this second service has shown that our objective of creating a platform for hosting communication services for deaf people has been successfully reached. SUMMA061 is currenty being used in the Medical Emergency Call Center of Madrid.

Due to the characteristics of the design that have been introduced, AIS and SUMMA061 are good frameworks for developing new communication services by creating the adequate modules, and maybe defining some new messages or events. No further changes would be needed.

We expect that in a short period of time deaf people enjoy new possibilities of communication, finding them useful. Furthermore, we are working on getting a complete communication between deaf and non deaf people in a future. However, difficult tasks involving speech recognition and synthesis (among others) must be faced in the way.

References

1. J.A. Cerrada, M. Collado. "Introduccin a la Ingeniera Software". UNED, 1995.
2. GlobalCall ISDN Technology Users Guide for Unix and Windows NT, Dialogic, 1998.
3. MSDN Library-October 1998 Release. Microsoft Corporation. 1995-1998.
4. Jernimo Arenas-Garca "Sistema de comunicacin en modo texto entre telfonos mviles Nokia 9110 y telfonos fijos EDT para personas sordas". (PFC). Madrid, 2000.
5. Survey of Text Telephones and Relay Services in Europe , EUR 14242 EN, CEC, Luxembourg 1992, ISBN 92-826-4395-6.

Captioning for Deaf and Hard of Hearing People by Editing Automatic Speech Recognition in Real Time

Mike Wald

Learning Technologies Group, School of Electronics and Computer Science
University of Southampton, United Kingdom
M.Wald@soton.ac.uk

Abstract. Deaf and hard of hearing people can find it difficult to follow speech through hearing alone or to take notes when lip-reading or watching a sign-language interpreter. Notetakers summarise what is being said while qualified sign language interpreters with a good understanding of the relevant higher education subject content are in very scarce supply. Real time captioning/transcription is not normally available in UK higher education because of the shortage of real time stenographers. Lectures can be digitally recorded and replayed to provide multimedia revision material for students who attended the class and a substitute learning experience for students unable to attend. Automatic Speech Recognition can provide real time captioning directly from lecturers' speech in classrooms but it is difficult to obtain accuracy comparable to stenography. This paper describes the development of a system that enables editors to correct errors in the captions as they are created by Automatic Speech Recognition.

1 Introduction

Deaf and hard of hearing people can find it difficult to follow speech through hearing alone or to take notes while they are lip-reading or watching a sign-language interpreter. Notetakers can only record a small fraction of what is being said while qualified sign language interpreters with a good understanding of the relevant higher education subject content are in very scarce supply [21]. UK Disability Discrimination Legislation states that reasonable adjustments should be made to ensure that disabled students are not disadvantaged [22]. Many systems have been developed to digitally record and replay multimedia face to face lecture content to provide revision material for students who attended the class or to provide a substitute learning experience for students unable to attend the lecture [1,5] and a growing number of universities are supporting the downloading of recorded lectures onto students' iPods or MP3 players [30]. As video and speech become more common components of online learning materials, the need for captioned multimedia with synchronised speech and text, as recommended by the Web Accessibility Guidelines [31], can be expected to increase and so finding an affordable method of captioning will become more important. Automatic Speech Recognition (ASR) can be used to create synchronised

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 683 – 690, 2006.

captions for live speech and multimedia material [4] and this paper will discuss methods to overcome existing problems with the technology by editing in real time to correct errors.

2 Use of Captions and Transcription in Education

Stinson [27] reported that deaf or hard of hearing students at Rochester Institute of Technology who had good reading and writing proficiency preferred real-time verbatim transcribed text displays to interpreting and/or notetaking. An experienced trained 're-voicer' using ASR by repeating very carefully and clearly what has been said can improve accuracy over the original speaker using ASR where the original speech is not of sufficient volume or quality or when the system is not trained. Re-voiced ASR is sometimes used for live television subtitling in the UK [14] as well as in courtrooms and classrooms in the US [8] using a mask to reduce background noise and disturbance to others. The most accurate system is real time captioning using stenographers and a special phonetic keyboard but trained stenographers in the UK do not choose to work in universities rather than in court reporting. Downs [7] noted the potential of speech recognition in comparison to summary transcription services and students in court reporting programs unable to keep up with the information flow in the classroom. Robison [20] identified the value of Speech Recognition to overcome the difficulties sign language interpreting had with foreign languages and specialist subject vocabulary for which there are no signs. Automatic speech recognition offers the potential to provide automatic real-time verbatim captioning archived as accessible lecture notes for deaf and hard of hearing students who may find it easier to follow the captions and transcript than to follow the speech of the lecturer.

3 ASR and Liberated Learning Concept

Feasibility trials using existing commercially available ASR software to provide a real time verbatim displayed transcript in lectures for deaf students in 1998 by the author in the UK [32] and St Mary's University, Nova Scotia in Canada identified that standard speech recognition software (e.g. Dragon, ViaVoice [18]) was unsuitable; it required the dictation of punctuation, which does not occur naturally in spontaneous speech in lectures. The international Liberated Learning Collaboration was established by Saint Mary's University, Nova Scotia, Canada in 1999 and since then the author has continued to work with IBM and Liberated Learning to investigate how ASR can make speech more accessible. Further investigations demonstrated the possibility of developing an ASR application that automatically formatted the transcription by breaking up the continuous stream of text based on the length of silences in the speech to provide a visual indication of pauses. The potential of using ASR to provide automatic captioning of speech in higher education classrooms has now been demonstrated in 'Liberated Learning' classrooms in the US, Canada and Australia [4,15,33]. Lecturers spend time developing their ASR voice profile by training the ASR software to understand the way they speak. They wear wireless microphones providing the freedom to move around as they are talking, while the text is displayed in real time on a screen using a data projector so students can simultaneously see and hear the lecture as it is delivered. After the lecture the text is edited for errors and made available for students on the Internet. To make the Liberated Learning vision a reality, the prototype ASR application, Lecturer developed in 2000 in collaboration with IBM was superseded the following year by IBM ViaScribe [11,3]. Both applications used the ViaVoice ASR 'engine' and its corresponding training of voice and language models and automatically provided text displayed in a window and stored for later reference synchronised with the speech. ViaScribe created files that enabled synchronised audio and the corresponding text transcript and slides to be viewed on an Internet browser or through media players that support the SMIL 2.0 standard [24] for accessible multimedia. ViaScribe can automatically produce a synchronised captioned transcription of spontaneous speech using automatically triggered formatting from live lectures, or in the office, or from recorded speech files on a website.

4 Improving Accuracy Through Editing in Real Time

Detailed feedback [15] from students with a wide range of physical, sensory and cognitive disabilities and interviews with lecturers showed that both students and teachers generally liked the Liberated Learning concept and felt it improved teaching and learning as long as the text was reasonably accurate (e.g. >85%). While it has proved difficult to obtain an accuracy of over 85% in all higher education classroom environments directly from the speech of all teachers, many students developed strategies to cope with errors in the text and the majority of students used the text as an additional resource to verify and clarify what they heard. Editing the synchronised transcript after a lecture, involving frequent pausing and replaying sections of the recording, can take over twice as long as the original recording for 15% error rates while for high error rates of 35%, it can take as long as if an audio typist had just completely transcribed the audio recording [3]. The methods used for enabling real time editing to occur can equally be applied to speed up post lecture editing and make it more efficient. Although it can be expected that developments in ASR will continue to improve accuracy rates [9,10,19], the use of a human intermediary to improve accuracy through correcting mistakes in real-time as they are made by the ASR software could, where necessary, help compensate for some of ASR's current limitations. Since not all errors are equally important, the editor can use their knowledge and experience to prioritise those that most affect readability and understanding. Lambourne [14] reported that although their ASR television subtitling system was designed for use by two operators, one revoicing and one correcting, an experienced speaker could achieve recognition rates without correction that were acceptable for live broadcasts of sports such as golf. Previous research has found that although ASR can transcribe at normal rates of speaking, correction of errors is problematic. Bailey [2] has reported that people type on computers typically between 20 and 40 words per minute. Lewis [16] found that corrections were made three times faster with voice, keyboard, and mouse than voice-only corrections. Karat [12] found that correction took over three times as long as entry time for experienced ASR users with good typing skills. Karat [13] found that novice users have similar numbers of speech and typing errors, but take much longer to correct dictation errors than typing errors whereas experienced users of ASR preferred keyboard-mouse techniques rather than speech-based techniques for making error corrections. Suhm [28] reported that speech recognition correction methods using spelling/handwriting/pen 'gesture' were of particular value for small mobile devices or users with poor typing skills. Shneiderman [23] noted that using a mouse and keyboard for editing required less mental effort than using speech.

5 Methods of Real Time Editing

Correcting ASR errors requires the editor(s) to notice that an error has occurred; moving a cursor into the position required to correct the error(s); typing the correction; continuing to listen and remember what is being said while searching for and correcting the error. There are many potential approaches and interfaces for real time editing, and these are being investigated to compare their benefits and to identify the knowledge, skills and training required of editors. Using the mouse and keyboard might appear the most natural method of error correction, but using the keyboard only for both navigation and correction and not the mouse has the advantage of not slowing down the correction process by requiring the editor to take their fingers off the keyboard to move the mouse to navigate to the error and then requiring the hand using the mouse to return to the keyboard for typing the correction. The use of foot operated switches or 'foot pedals' to select the error and using the keyboard to correct the error has the advantage of allowing the hands to concentrate on correction and the feet on navigation, a tried and tested method used by audio typists [26]. Separating the tasks of selection and correction and making correction the only keyboard task, has the advantage of allowing the editor to begin typing the correct word(s) even before the error selection has been made using the foot pedal. An ASR editing system that separated out the tasks of typing in the correct word and moving the cursor to the correct position to correct the error would facilitate the use of two editors. As soon as one editor spotted an error they could type the correction and these corrections could go into a correction window. The other editor's role would be to move a cursor to the correct position to replace the error with the correction. For low error rates one editor could undertake both tasks. An alternative approach is for errors to be selected sequentially using the tab key or foot switch or, through random access by using a table/grid where selection of the words occurs by row and column position. If eight columns were used corresponding to the 'home' keys on the keyboard and rows were selected through multiple key presses on the appropriate column home key, the editor could keep their fingers on the home keys while navigating to the error, before typing the correction. Real time television subtitling has also been implemented using two typists working together to overcome the difficulties involved in training and obtaining stenographers who use a phonetic keyboard or syllabic keyboards [25,17]. The two typists can develop an understanding to be able to transcribe alternate sentences, however only stenography using phonetic keyboards is capable of real time verbatim transcription at speeds of 240 words per minute. For errors that are repeated (e.g. names not in the ASR dictionary) corrections can be offered by the system to the editor, with the option for the editor to accept or reject them. Although it is possible to devise 'hot keys' to 'automatically' correct some errors (e.g. plurals, possessives, tenses, a/the etc.) the cognitive load of remembering the function of each key may make it easier to actually correct the error directly through typing. Speech can be used to correct the error, although this introduces another potential error if the speech is not recognised correctly. Using the speech to navigate to the error by speaking the coordinates of the error is a possibility, although again this would involve verbal processing and could overload the editor's cognitive processing as it would give them even more to think about and remember.

6 Feasibility Test Methods, Results and Evaluation

A prototype real-time editing system with editing interfaces using the mouse and keyboard, keyboard only and keyboard only with the table/grid was developed to investigate the most efficient approach to real-time editing. Five test subjects were used who varied in their occupation, general experience using and navigating a range of software, typing skills, proof reading experience, technical knowledge about the editing system being used, experience of having transcribed speech into text, and experience of audio typing. Different 2 minute samples of speech were used in a randomised order with speech rates varying from 105 words per minute to 176 words per minute and error rates varying from 13% to 29%. Subjects were tested on each of the editing interfaces in a randomised order, each interface being used with four randomised 2 minutes of speech, the first of which was used to give the user practice to get used to how each editor functioned. Each subject was tested individually using a headphone to listen to the speech in their own quiet environment. In addition to quantitative data recorded by logging, subjects were interviewed and ask to rate each editor. Navigation using the mouse was preferred and produced the highest correction rates. However this study did not use expert typists trained to the system who might prefer using only the keyboard and obtain even higher correction rates. An analysis of the results showed there appeared to be some learning effect suggesting that continuing to practice with an editor might improve performance. All 5 subjects believed the task of editing transcription errors in real-time to be feasible and the objective results support this as up to 11 errors per minute could be corrected, even with the limited time available to learn how to use the editors, the limitations of the prototype interfaces and the cognitive load of having to learn to use different editors in a very short time.

7 Automatic Error Correction

Future work will include investigating automatic error correction using phonetic searching and confidence scores which reflect the probability that the recognised word is correct. Suhm [29] found that highlighting likely errors based on these confidence scores didn't help speed up correction as some correct words were also highlighted as well. Phonetic searching [6] can help find ASR 'out of vocabulary' errors that occur when words spoken are not known to the ASR system, as it searches for words based on their phonetic sounds not their spelling. If the system can compare the phonemes of the correct word typed in by the editor with the phonemes of the possible words that contain the error, then coupled with the confidence scores it may be possible to automatically identify the error and replace it with the typed correction, with the option for the editor of overriding the automatic system if it makes a mistake. The system could begin to compare the phonemes of the correct word as it is typed in even before the whole word has been entered.

8 Methods of Displaying the Edited Text

It is possible to display both errors and corrections on the large screen as they are being corrected. Liberated Learning's research has shown that while projecting the text onto a large screen in the classroom has been used successfully it is clear that in many situations an individual personalised and customisable display would be preferable or essential. If text is displayed only after editing a variable delay would be introduced as errors can be bunched together. A client server personal display system has been developed [34] to provide users with their own personal display on their own wireless systems customised to their display preferences. This also enables the ASR transcriptions of multiple speakers using multiple ASR systems to be displayed on multiple personal display windows. These individual display 'captions' could be combined into one window with speakers identified to provide a corrected live transcript of a meeting. A deaf person may then be able to cope even better than the hearing listeners who would struggle with the auditory interference of everyone talking at once.

9 Conclusion

Understanding speech and notetaking in lectures can be difficult for deaf students and so ASR could be very helpful to them. Improving the accuracy of the live ASR captions and transcript and developing faster editing methods is important because editing can be difficult and slow. Some ASR errors may have a negligible effect on readability and knowledge of this would enable editors to prioritise error correction for those errors most affecting readability if they were unable to correct 100% of the errors. The optimal system to digitally record and replay multimedia face to face lecture content would automatically create an error free transcript of spoken language synchronised with audio, video, and any on-screen graphical display (e.g. PowerPoint) and enable this to be displayed in different ways in different devices. Real time editing was shown to be feasible, but the relatively small subject and test sample size, the lack of the test subjects' practice with the editing interfaces and the high cognitive load of having to change to a new and different editor approximately every 20 minutes meant that the results, while indicative, are not conclusive but can be helpful in informing the direction of future developments. Continued research is therefore needed to improve the accuracy of ASR and develop an efficient method of editing in real time before the Liberated Learning vision can become an everyday reality.

References

1. Baecker, R. M., Wolf, P., Rankin, K. The ePresence Interactive Webcasting System: Technology Overview and Current Research Issues. *Proceedings of Elearn 2004* (2004) 2396-3069
2. Bailey. Human Interaction Speeds. Retrieved December 8, 2005, from `http://webusability.com/article_human_interaction_speeds_9_2000.htm`

3. Bain, K., Basson, S., A., Faisman, A., Kanevsky, D. Accessibility, transcription, and access everywhere, IBM Systems Journal, Vol 44, no 3, (2005) pp. 589-603 Retrieved December 12, 2005, from http://www.research.ibm.com/journal/sj/443/bain.pdf
4. Bain, K., Basson, S., Wald, M. Speech recognition in university classrooms, Proceedings of the Fifth International ACM SIGCAPH Conference on Assistive Technologies, ACM Press, (2002) 192-196.
5. Brotherton, J. A., Abowd, G. D. Lessons Learned From eClass: Assessing Automated Capture and Access in the Classroom, ACM Transactions on Computer-Human Interaction, Vol. 11, No. 2. (2004)
6. Clements, M., Robertson, S., Miller, M. S. Phonetic Searching Applied to On-Line Distance Learning Modules (2002). Retrieved December 8, 2005, from http://www.imtc.gatech.edu/news/multimedia/spe2002_paper.pdf
7. Downs, S., Davis, C., Thomas, C., Colwell, J. Evaluating Speech-to-Text Communication Access Providers: A Quality Assurance Issue, PEPNet 2002: Diverse Voices, One Goal Proceedings from Biennial Conference on Postsecondary Education for Persons who are Deaf or Hard of Hearing. (2002) Retrieved November 8, 2005, from http://sunsite.utk.edu/cod/pec/2002/downs.pdf
8. Francis, P.M. Stinson, M. The C-Print Speech-to-Text System for Communication Access and Learning, Proceedings of CSUN Conference Technology and Persons with Disabilities, California State University Northridge. (2003) Retrieved December 12, 2005, from http://www.csun.edu/cod/conf/2003/proceedings/157.htm
9. Howard-Spink, S. IBM's Superhuman Speech initiative clears conversational confusion. (2005). Retrieved December 12, 2005, from http://www.research.ibm.com/thinkresearch/pages/2002/20020918_speech.shtml
10. IBM. The Superhuman Speech Recognition Project (2003) Retrieved December 12, 2005, from http://www.research.ibm.com/superhuman/superhuman.htm
11. IBM (2005) Retrieved December 12, 2005, from http://www-306.ibm.com/able/solution_offerings/ViaScribe.html
12. Karat, C.M., Halverson, C., Horn, D. and Karat, J. Patterns of Entry and Correction in Large Vocabulary Continuous Speech Recognition Systems, CHI 99 Conference Proceedings, (1999) 568-575.
13. Karat, J., Horn, D., Halverson, C. A., Karat, C.M. Overcoming unusability: developing efficient strategies in speech recognition systems, Conference on Human Factors in Computing Systems CHI '00 extended abstracts, (2000) 141-142.
14. Lambourne, A., Hewitt, J., Lyon, C., Warren, S. Speech-Based Real-Time Subtitling Service, International Journal of Speech Technology, 7, (2004) 269-279.
15. Leitch, D., MacMillan, T. Liberated Learning Initiative Innovative Technology and Inclusion: Current Issues and Future Directions for Liberated Learning Research. Year III Report. (2003) Saint Mary's University, Nova Scotia.
16. Lewis, J.R. Effect of Error Correction Strategy on Speech Dictation Throughput, Proceedings of the Human Factors and Ergonomics Society, (1999) 457-461
17. NCAM International Captioning Project (2000) Retrieved December 12, 2005, from http://ncam.wgbh.org/resources/icr/europe.html
18. Nuance (2005) Retrieved December 12, 2005, from http://www.nuance.com/
19. Olavsrud, T. (2002). IBM Wants You to Talk to Your Devices Retrieved December 12, 2005, from http://www.internetnews.com/ent-news/article.php/1004901
20. Robison, J., Jensema, C. Computer Speech Recognition as an Assistive Device for Deaf and Hard of Hearing People, Challenge of Change: Beyond the Horizon, Proceedings from

Seventh Biennial Conference on Postsecondary Education for Persons who are Deaf or Hard of Hearing. April, 1996. (1996) Retrieved November 8, 2005, from http://sunsite.utk.edu/cod/pec/1996/robison.pdf
21. RNID (2005). Retrieved December 12, 2005, from http://www.rnid.org.uk/howwehelp/research_and_technology/communication_and_broadcasting/virtual_signing/
22. SENDA (2001). Retrieved December 12, 2005, from http://www.opsi.gov.uk/acts/acts2001/20010010.htm
23. Shneiderman, B. The Limits of Speech Recognition, Communications of the ACM September 2000, Vol. 43(9), (2000) 63-65
24. SMIL (2005) Retrieved December 12, 2005, from http://www.w3.org/AudioVideo/
25. Softel. FAQ Live or 'Real-time' Subtitling (2001) Retrieved December 12, 2005, from http://www.softel-usa.com/downloads/Softel_Live_Subtitling_FAQ.pdf
26. Start-Stop Dictation and Transcription Systems (2005). Retrieved December 27, 2005, from http://www.startstop.com/sst2.asp
27. Stinson. M., Stuckless, E., Henderson, J., Miller, L. Perceptions of Hearing-Impaired College Students towards real-time speech to print: Real time Graphic display and other educational support services, The Volta Review, 90, (1988) 341-47
28. Suhm, B., Myers, B., Waibel, A. Model-Based and Empirical Evaluation of Multimodal Interactive Error Correction, CHI 99 Conference Proceedings, (1999) 584-591
29. Suhm, B., Myers, B. Multimodal error correction for speech user interfaces, ACM Transactions on Computer-Human Interaction (TOCHI), Vol. 8(1), (2001) 60-98
30. Tyre, P. Professor In Your Pocket, Newsweek MSNBC (2005) Retrieved December 8, 2005, from http://www.msnbc.msn.com/id/10117475/site/newsweek
31. WAI. (2005). Retrieved December 12, 2005, from http://www.w3.org/WAI
32. Wald, M. Developments in technology to increase access to education for deaf and hard of hearing students, Proceedings of CSUN Conference Technology and Persons with Disabilities, (2000) California State University, Northridge. Retrieved December 12, 2005, from http://www.csun.edu/cod/conf/2000/proceedings/0218Wald.htm
33. Wald, M. Hearing disability and technology, Phipps, L., Sutherland, A., Seale, J. (Eds) Access All Areas: disability, technology and learning, JISC TechDis and ALT, (2002) 19-23.
34. Wald, M. (2005) Personalised Displays. Speech Technologies: Captioning, Transcription and Beyond, IBM T.J. Watson Research Center New York Retrieved December 27, 2005, from http://www.nynj.avios.org/Proceedings.htm

Improving Baby Caring with Automatic Infant Cry Recognition

Sandra E. Barajas-Montiel[1], Carlos A. Reyes-García[1], Emilio Arch-Tirado[2], and Mario Mandujano[2,3]

[1] Instituto Nacional de Astrofisica Optica y Electronica, Luis Enrique Erro #1, Tonantzintla, Puebla, Mexico
{sandybarajas, kargaxxi}@ccc.inaoep.mx
[2] Instituto Nacional de Rehabilitación
earch@inr.gov.mx
[3] UAM-Xochimilco
mariom@att.net.mx

Abstract. Babies are human beings who cannot satisfy their necessities by themselves, they completely depend of cares and attentions by adults. The cry is the natural media babies use to express their needs. Several studies have demonstrated that cry is a useful tool to determine the different emotional and physiological states from an infant, and in addition to make medical diagnoses of diseases related to the central nervous system. This work presents the analysis and extraction of characteristics from infant crying for its automatic classification with Support Vector Machines. Several classification tasks were done, working in the identification of pain, hunger, and deafness levels with results of up to 96 % of correct classification. Besides some results, we show the implementation and experimentation done.

1 Introduction

Babies express their needs and feelings through cry. Several researches have demonstrated that infant cry is a powerful tool that can be used to decipher the needs of the babies or for making medical diagnoses of pathologies at very early stages of life. Those studies open the opportunity of helping babies by means of the information contained in their crying. There's a tight relationship between the central nervous system and the efferent ways that involve the phono-articulator apparatus. When this function is not the proper one, changes are produced in a intrinsic laryngeal muscular level so as in the muscles that intervene in the respiratory process, these changes will translate possibly in alterations or modifications in the baby's cry. If we are able to understand better what they need, in their pre-speech stage of life, it would be easier to take care of them. Moreover, if we were able to detect certain pathologies we could opportunely provide the appropriate treatment and prevent future complications and undesirable sequels. World-wide statistics indicate that from one thousand new born, 1 or 2 may have a perceptive difficulty, presenting deep or severe hearing loss. Children with normal cognitive development whose hearing losses are identified before six

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 691 – 698, 2006.

months can develop language at the same or a similar rate to a hearing child [16]. Children identified with a hearing loss between birth and six months old have a receptive language of 200 words and expressive language of 117 words, whereas those identified between ages of seven and 18 months have a receptive language of 86 words and expressive language of 54 words. When tested at 26 months of age, those identified as deaf before six months old have "significantly higher" measures of language growth and personal-social development.

The first works with infant cry were initiated by Wasz-Hockert since the beginnings of the 60's. In 1964 Wasz-Hockert showed that the four basic types of cry can be identified by listening: pain, hunger, pleasure and birth [1]. Since then, many other studies related to this line of research have been reported. Sergio D. Cano has carried out and directed several works devoted to the extraction and automatic classification of acoustic characteristics of infant cry. In one of those studies, in 1999 Cano presented a work in which he demonstrates the utility of Self-Organizing Maps in the classification of Infant Cry Units [2]. In [3] a radial basis function (RBF) network is implemented for infant cry classification in order to find out relevant aspects concerned with the presence of CNS diseases. The cry units are classified in 2 categories, normal or abnormal. Reyes and Orozco [4] classified cry samples from deaf and normal babies with feed-forward neural networks, reporting a 97.43% of efficiency. In [5] it is showed the implementation of a Fuzzy Relational Neural Network for Detecting Pathologies by Infant Cry Recognition with percentage of correct classification of 97.3% and 98%. Petroni, Malowany, Johnston and Stevens [6] classify cry from normal babies to identify pain with different artificial neural networks. They reported from 61% up to 86.2% of precision in the classification.

Among the models for pattern classification, neural networks have been used to extract patterns and detect trends that are too complex to be noticed by either humans or other computer techniques. Another recently becoming popular classification technique is the Support Vector Machine (SVM), which was developed by Vapnik and his group at AT&T Bell laboratories [7]. SVM's have been used for isolated handwritten digit recognition, object recognition, speaker identification, charmed quark detection, face detection in images, and text categorization. Experimental results indicate that SVM's can achieve a generalization performance that is greater than or equal to other classifiers, while requiring significantly less training data to achieve good results [8]. In this paper we present the result of the use of SVM's for infant cry recognition to detect hunger and pain and to differentiate deafness levels.

2 The Automatic Infant Cry Recognition Process

Automatic Infant Cry Recognition (AICR) is very similar to the Speech Recognition Process. Basically the AICR is performed in two stages. The first section is for signal processing and the second one is for pattern classification. In the signal processing phase, the cry signal is first normalized and cleaned, and then it is analyzed to extract the most important characteristics in function of time. In AICR like in any pattern recognition problem, the goal is that given an input pattern we obtain as an output at the end of the recognition process the class to which this pattern belongs.

2.1 Signal Processing Phase

The acoustical analysis of the raw cry wave form provides the information needed for its recognition. At the same time, it discards unwanted information such as background noise and channel distortion [9]. In this phase we make a transformation of measured data into pattern data. There are several techniques for analyzing cry wave signals, some of them are: Linear Prediction Coefficient (LPC), Mel Frequency Cepstral Coefficients (MFCC), Intensity, and Spectral Analysis. For our experiments we use Mel Frequency Cepstral Coefficients. MFCC are analog to perceptual characteristics. They can be obtained like filtered signals through different frequency scales. The Mel spectrum operates on the basis of selective weighting of the frequencies in the power spectrum. High order frequencies are weighted on a logarithmic scale whereas lower order frequencies are weighted on a linear scale. This technique pretends to simulate the properties the ear has as a filter, which is more sensitive to some frequencies than to others [10,11].

These coefficients are calculated in small frames of the signal on time. Only the first *M* cepstral coefficients are used as features. The spectral form is modelled by the first coefficients and the precision depends on the number of coefficients taken. In Fig.1 we can see the general scheme to obtain MFCC. The applications of this technique during signal processing results in the production of values of a set of acoustic features. The set of values for *n* features may be represented by a vector in an *n*-dimensional space. Each vector represents a pattern.

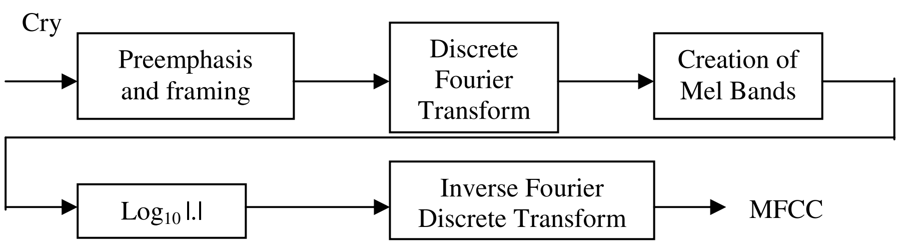


Fig. 1. Scheme to obtain MFCC

2.2 Pattern Recognition Phase

The Infant Cry Recognition Process is a binary task where the classifier decides whether or not a cry belongs indeed to the class it claims to be. The set of *n* feature vectors is divided in two subsets: the training set and the test set. First the training set is used to teach the classifier to distinguish between the different cry types. Then the test set is used to determine how well the classifier assigns the corresponding class to a pattern by means of a prediction rule generated during training. Fig. 2 shows the general process of the pattern recognition phase.

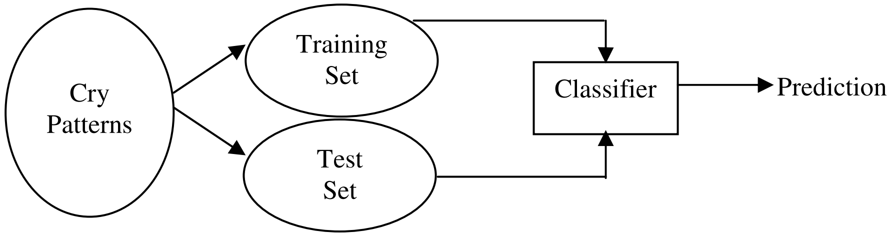


Fig. 2. Pattern Classification Phase

3 System Implementation

In the present work we'll focus on the description of Support Vector Machines. We have selected this kind of model, in principle, because of their adaptating and learning capacity. SVM's have been gaining wide acceptance because of their high generalization ability for many kinds of applications.

3.1 Support Vector Machines

The algorithm of Support Vector Machines (SVMs) is an alternative training technique to Polynomial, Radial Basis Function and Multi-Layer Perceptron classifiers [12]. Support Vector Machines are based on the Structural Risk Minimization principle [13] from computational learning theory. An SVM is a binary classifier that makes its decisions by constructing a linear decision boundary or hyperplane that optimally separates two classes. Fig. 3 shows an example of the optimal hyperplane for a problem. It uses the discriminant function $f{:}X \subseteq \mathrm{R}^n \rightarrow \mathrm{R}$ of the following form:

$$f(x) = \langle \alpha \cdot k_s(x) \rangle + b. \tag{1}$$

The $k_S(x) = [k(x, s_1), \ldots, k(x, s_d)]^T$ is the vector of evaluations of kernel functions centered at the support vectors $S = \{s_1, \ldots, s_d\}$, $s_i \in \mathrm{R}^n$ which are usually subsets of the training data. The $\alpha \in R^l$ is a weight vector and $b \in R$ is a bias. The linear SVM aims to train the linear discriminant function:

$$f(x) = \langle w \cdot x \rangle + b. \tag{2}$$

of the binary classifier [14]:

$$q(x) = \begin{cases} 1 \; for \; f(x) \geq 0, \\ 2 \; for \; f(x) < 0. \end{cases} \tag{3}$$

For linearly separable data labeled $\{x_i, y_i\}$, $x_i \in R^d$, $y_i \in \{-1,1\}$, $i = 1 \ldots N$ the optimal hyperplane is chosen according to the maximum margin criterion, i.e. by choosing the separating plane that maximizes the Euclidean distance to the nearest data points on each side of that plane. This is achieved by minimizing the square of the L2-norm of w, $\|w\|_2^2$ coupled to the inequalities:

$$(x \cdot w + b) y_i \geq 1 \;\; for \;\; all \;\; i. \tag{4}$$

The solution for the optimal hyperplane, w_0, is a linear combination of a small subset of data, x_s, $s \in \{1 \ldots N\}_g$, known as the support vectors. These support vectors also satisfy the equality

$$(\ (x_s \cdot w_0 + b)\ ) y_s = 1 \tag{5}$$

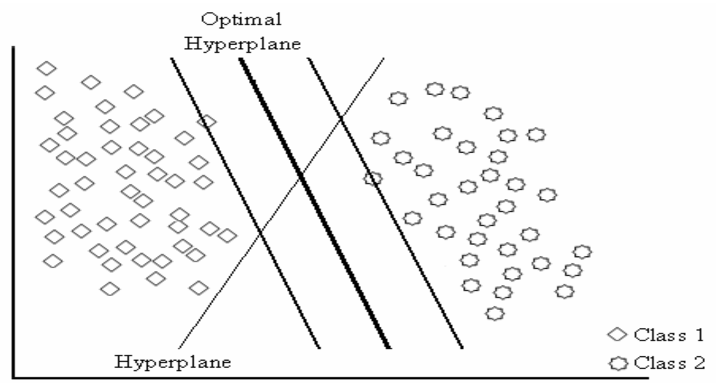


Fig. 3. An example of the Optimal hyperplane

3.2 Experimentation

For the present experiments we worked with samples of infant cries. The infant cries were collected by recordings done directly by medical doctors, the cry samples were carefully labeled at the time of the recording with references like infant age and the reason for the crying. Then each signal wave was divided in 1 second segments, each segment represents a sample. Next, MFCC acoustic features were obtained by means of the freeware program Praat v4.0.8 [15]. Every 1 second sample is divided in 50-millisecond frames from which 16 coefficients are extracted, this procedure generates vectors with 304 coefficients each. In order to reduce the sample vectors dimensions we apply Principal Component Analysis (PCA).

We used a corpus composed by 209 pain cry samples, and 1418 no-pain samples, these samples include hunger, sleepy and uncomfortable types. We also used a corpus with 759 hunger cry samples, and 868 no-hunger samples. Moreover, we worked with a corpus of 531 samples of the Medium Hearing Loss (MHL) class and 65 of the class Moderately Severe Hearing Loss. (MSHL).

The PCA algorithm was implemented in Matlab 6.5. For the implementation of Support Vector Machines we used the Matlab's Stprtool toolbox [14]. Support Vector Machines use the Radial Basis Function as kernel. All the parameter values of the classifier were established in a heuristic way after performing several experiments. During the experiments we used the 10-fold cross validation technique to evaluate the performance and reliability of the classifier.

3.3 Preliminary Results

Three different classification experiments were performed, the first one consists in classifying pain and no pain cry (P- NP), the second one was made to classify infant cry in categories called hunger and no hunger (H-NH), and the third one was about the identification of deafness levels (MHL-MSHL). The results of these classification experiments are shown in Table1.

The classification accuracy was calculated by taking the number of samples correctly classified, divided by the total number of samples. Different number of principal components (PC's) was used at each classification task. The number of PC's tested in the experiments showed here was 2, 3, 10, 16 and 50.

The results are the average of several different experiments. Comparing the results showed here with the results reported in reference [6], in which also pain is identified, we can notice better performance when SVM are used. The results suggest that the application of MFCC is a good option for automatic infant cry recognition. They allow the system to differentiate emotional status of the babies and also provides a guide for medical diagnoses in hearing disorders.

Table 1. Results of the three classification experiments

Problem	% Classification Accuracy				
	PCA2	PCA3	PCA10	PCA16	PCA50
P-NP	88	83.333	93.333	96	81.667
H-NH	57.333	48.667	73.333	78	77.333
MHL-MSHL	91.333	91.333	87.333	91.333	88.667

4 Conclusions and Future Work

This work, as other studies, shows that there is much information inside the infant cry waves, which can be used to improve the care of recently born infants. Furthermore, the most important thing is that this information may be used to support medical diagnosis and to warn pediatricians of possible present pathologies, avoiding the use of other more invasive and uncomfortable techniques for babies and their parents. In this way, doctors can use automatic infant cry recognition for detecting pathologies at early stages of live and provide the suitable treatment and prevent future complications in the infant's life.

In this paper we present the automatic classification of infant cry done by means of Support Vector Machines. The best results were obtained in the classification of Pain and No Pain cries reaching 96 % of correct classification. For the problem H-NH the highest classification accuracy obtained was 78 % and for the MHL.MSHL classification task the best correct classification was 91.33 %. For all the performed experiments, 16 was the best number of Principal Components. As we have shown, if we use more or less than 16 PCs the classification accuracy decreases, while, in the first case, the computation time increases.

The correct classification results obtained until this moment are very encouraging. With a larger number of samples we could generalize better our results, in order to be closer to end up with a real time practical system, ready to be applicable by any person in any infant's caring environment.

In this work we used the MFCC acoustical features, but there are other alternatives that can be useful like; fundamental frequencies, cochleograms, linear prediction cepstral coefficients, or the combination of several acoustic features in the same pattern. On the basis of new approaches for speech recognition, other models or other types of ensembles also can be tested. We are in the process of testing several kinds of neural networks, and also testing several kinds of hybrid models, like combining neural networks with genetic algorithms and fuzzy logic, or other complementary models. The collection of more samples will also allow us to have a larger number of classes for the classification of normal cry and also for the classification of more deafness levels. It is difficult to find the type of cry we require at the precise time, but with enough samples a classification of cry can be applied to detect pain, hunger, etc from deaf babies too. And finally, we would like to collect the adequate sets of samples directed to allow us the identification of as many pathologies as possible.

Acknowledgments

This work is part of a project that is being financed by CONACYT-Mexico (C01-46753).

References

1. Wasz-Hockert, O., Partanen, T., Vuorenkoski, V., Valanne, E. and Michelsson, K, "The Identification of Some Specific Meanings in Infant Vocalization," Experiencia, vol. 20, pp. 154-156, 1964.
2. S.D. Cano-Ortiz, D. I. Escobedo-Becerro, "Clasificación de Unidades de Llanto Infantil Mediante el Mapa Auto-Organizado de Koheen," I Taller AIRENE sobre Reconocimiento de Patrones con Redes Neuronales, Universidad Católica del Norte, Chile,pp. 24-29, 1999.
3. Sergio D. Cano Ortiz, Daniel I. Escobedo Beceiro, Taco Ekkel2, "A Radial Basis Function Network Oriented for Infant Cry Classification", Proc. Of 9th Iberoamerican Congress on Pattern Recognition, Puebla, Mexico, 2004.
4. J., Reyes, C.A., Reyes, "Mel-frequency Cepstrum Coefficients Extraction from Infant Cry for Classification of Normal and Pathological Cry whit Feed-Forward Neural Networks," Proc. of ESANN, Bruges, Belgium, 2003.
5. Suaste, I., Reyes, O.F., Diaz, A., Reyes, C.A., "Implementation of a Linguistic Fuzzy Relational Neural Network for Detecting Pathologies by Infant Cry Recognition," Proc. of IBERAMIA, Puebla, Mexico , pp. 953-962, 2004.
6. M. Petroni, A.S. Malowany, C.C. Johnston, B.J. Stevens, "Identification of Pain from Infant Cry Vocalizations Using Artificial Neural Networks," Proc. of SPIE, vol. 2492, pp. 729-738, 1995.
7. C. Cortes and v. Vapnik, "Support vector networks", Machine Learning, vol. 20, pp. 1-25, 1995.

8. V. Wan, W.M. Campbell, “Support Vector Machines for Speaker Verification and Identification,” IEEE International Workshop on Neural Networks for Signal Processing, Sydney, Australia, 2000.
9. Livinson S.E., and Roe, D.B, “A Perspective on Speech Recognition,” IEEE Communications Magazine, pp. 28-34, 1990.
10. Orosco GJ, Reyes CA, “Mel-Frequency Cepstrum Coefficients Extraction from Infant Cry for Classification of Normal and Pathological Cry with Feed-Forward Neural Networks,” Proc. International Joint Conference on Neural Networks. Portland, Oregon, USA, pp. 3140-3145, 2003.
11. Reyes O, Reyes CA., “Clasificación de Llanto de Bebés para Identificación de Hipoacusia y Asfixia por medio de Redes Neuronales,” Proc. of the II Congreso Internacional de Informática y Computación de la ANIEL, Zacatecas, México, pp. 20-24, 2003.
12. E. Osun, R. Freund, F. Girosi, Support Vector Machines: Training and Applications, Massachusetts Institute of Technology, Cambridge, MA, USA, 1997.
13. Vladimir N. Vapnik, “The Nature of Statistical Learning Theory,” Springer, New York, 1995.
14. Vojtech Fran, Vaclav Hlavac, Statistical Pattern Recognition Toolbox, Czech Technical University Prague, 1999.
15. Boersma, P., Weenink, D., Praat v 4.0.8: A System for Doing Phonetics by Computer. Institute of Phonetic Sciences of the University of Amsterdam, 2002.
16. Yoshinaga-Itano 1998, quoted in *‘The High Cost Of Hearing Lost; What Our Publics Need to Know’,* Donald Radcliffe, The Hearing Journal, May 1998, vol 51 no. 5.

People with Cognitive Problems and the Aging Population
Introduction to the Special Thematic Session

Erwin Fugger, Martin Morandell, and Barbara Prazak

ARC Seibersdorf research GmbH, Department of Rehabilitation & Inclusion,
Viktor Kaplan Strasse 2, A-2700 Wr. Neustadt, Austria
{erwin.fugger, barbara.prazak, martin.morandell}@arcsmed.at

Abstract. As a result of increasing life expectancy and the resulting demographic shift, there are ever growing numbers of elderly persons, living in single households. In consequence of declining cognitive and physical abilities, many affected persons are increasingly dependent on external support to carry out basic and instrumental activities of everyday living. In order to extend the time span elderly people can live independently in their preferred environment, new solutions offered by modern information and communications technology (ICT) need to be introduced. In important respects, ICT support is the primary unsatisfied need of individual elderly persons, their families and caregivers. This need can be met by ICT–based support in health and activity monitoring, by enhancing safety and security in home environments, by improving access to social welfare and medical services, by facilitating social contacts as well as access to context-based infotainment.

1 Motivation

The generation 85+ is the most rapidly growing segment of the population in Western Europe and North America [1]. Accordingly, the predictable increasing number of elderly people and the related effects (especially the increase of age related diseases like Alzheimer-dementia) cause a need for new approaches in care giving. Supporting accessibility as well as strengthening inclusion is essential so as not to create socially isolated and disadvantaged individuals. By developing and utilising new instruments and assistive technologies (AT) for home environments, affected persons should be able to cope more easily with occurring cognitive disabilities. Consequently, AT should contribute to sustain a persons´ independent evening of life within a familiar environment, which is adaptable to changing needs.

One of the big challenges for this century is to cope with the expected ´Alzheimer Wave´. The number of people with Alzheimer will increase from 24.3 million today, to 42.3 million in 2020 and will double every 20 years. [2]. This is a consequence of the higher life expectancy and the shift of the baby-boom generation.

In this Special Thematic session a wide range of appropriate research projects in the proposed thematic fields will be presented. In all these projects two areas should be emphasized; this is on the one hand *Human Computer Interaction* and on the other hand *Ethical Issues*.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 699 – 700, 2006.

1.1 Human Computer Interaction (HCI)

If computer based devices and services are to be helpful for people not experienced in using technical devices as well as for people with limited cognitive abilities, new solutions for Human Computer Interaction have to be developed. Already a "double click" can be quite difficult for persons with limited eye-hand coordination. Simply expressed, computer interfaces have to become more human needs centred. If it, for example, would be possible to talk to a computer, the inhibition threshold to use such systems could be reduced considerably. Therefore also talking heads can play a significant role [3], giving context aware advices to the user.

When it is possible to create computer systems which are already easy to use for kids at pre-school age, it should also be possible to design appropriate interfaces for elderly persons. When designing a technical system for the elderly, one of the main challenges is, to allow the maximum amount of information transfer to users under conditions of minimum cognitive effort. [4] A standard "Text-Button-Icon" Graphical User Interface won't meet these requirements. Already the commonly used image metaphors applied in connection with icons won't be understood by people with dementia. To achieve a higher acceptance of new technical devices in the mentioned user groups, especially ´emotional computing´ will play an important role in the future.

2 Ethical Issues

The utilisation of ICTs in (home-) care settings has to regard especially ethical considerations respecting the dignity of all involved parties. Ending up in an ethical dilemma can occur when on the one hand values are tried to be kept (i.e. by not using monitoring systems), but on the other hand when it has the effect that other values are violated (i.e. independence) [5]. When talking about ´Computers and Caring´, it is often mentioned that the ´overuse´ of technical aids can lead to isolation of a user, to the replacement of helping hands. On the other side, it may contribute to become more independent from one's caretaker.

References

1. Whitehouse, P.: Can a computer be a caregiver? (2002) American Association for Artificial Intelligence
2. Cleusa, P., Ferri, M., e.a.: Global prevalence of dementia: a delphi consensus study. Lancet 366 (2005) 2112–7
3. Trenkwalder Andrea: Forscher wollen Verstaendigung zwischen Mensch und Maschine erleichtern. http://www.heise.de/newsticker/meldung/60270
4. Mihailidis, A., Carmichael, B., Boger, J.: The use of computer vision in an intelligent environment to support aging-in-place, safety, and independence in the home. IEEE Transactions on Information Technology in Biomedicine 8(3) (2004) 238–247
5. Riederer, E.: Explorative Entwicklung informationstechnischer Innovationen Systementwuerfe von Therapie- und Alltagshilfen fuer Menschen mit Demenz. PhD thesis, UNIVERSITAT DER BUNDESWEHR MUENCHEN (1999)

A Semi-autonomous Wheelchair Towards User-Centered Design

Sven Rönnbäck[1], Jouni Piekkari[2], Kalevi Hyyppä[1], Tomas Berglund[1], and Simo Koskinen[3]

[1] Department of Computer Science and Electrical Engineering
Luleå University of Technology SE-971 87 Luleå, Sweden
{Sven.Ronnback, Kalevi.Hyyppa, Tomas.Berglund}@ltu.se
[2] Faculty of Art and Design, University of Lapland
FIN-96101 Rovaniemi, Finland
Jouni.Piekkari@ulapland.fi
[3] Faculty of Social Science, University of Lapland
FIN-96101, Rovaniemi Finland
Simo.Koskinen@ulapland.fi

Abstract. Research on assistive technology for impaired and elderly is of great importance and it is intended to grow as society undergoes an age shift in its population. Research on assistive technology leads to the development of aids for individual users. These aids can be made more or less autonomous in order to fit an individuals specific needs. Aids can be designed not only to please a user from a technical perspective, but also from a psychological perspective. Based on knowledge about a user, from for example interviews, the design of an aid can be improved.

We present a semi-autonomous wheelchair which can be controlled using head-mounted sensors. Control is also possible by sensors placed on the hand of a user. The wheelchair hand control was tested by a user and the feedback from the user is included.

Through an interface suitable for specific users, the wheelchair can perform certain tasks autonomously. One such task is moving to a certain location pointed out by a user looking at a map of the surrounding which is presented on a computer screen.

With a user centered perspective based on interviews, direct contact, and knowledge about users, we show results for improving the design of assistive technology.[1]

1 Introduction

The Mobile Internet Connected Assistant (MICA) project [1,2] is a research project involving Luleå University of Technology in Sweden and University of Lapland in Finland.

At Luleå University, the research is focused on the technological aspects of assistive design, while at the University of Lapland the focus is on user centered design.

[1] This work was partly funded by EU and Interreg IIIA, Nord.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 701–708, 2006.

The main goal of the MICA project is to perform research aiming to improve the knowledge about assistive design that can be used to make a usable assistant or helper for a user. To reach this goal, recent knowledge and information about sensors, navigation methods, algorithms, steering equipment and other electrical equipment, is not sufficient but knowledge about the user is also important. As soon as assistive technology becomes available, there is currently a big interest in it [3,4].

A user centered design cares about the social and physical environment of a user. In the ELVI research project [5], a database of knowledge about users as well as their ages has been built. The reason for building this database was the pressure to improve design as we are approaching an age shift in the society and the population. The expected result from studying the database was information of possibilities on how the knowledge of users promotes the development of design practice according to current user needs.

Usually, a wheelchair is controlled with a joystick. By moving the joystick, the user controls the vehicle.

It is common to use a combination of manual control and autonomous control to take a wheelchair through a narrow passage and to reduce the risk of collision [6]. Other solutions can plan the motion of the vehicle to the middle of the detected free space defined by obstacles and the environmental layout or take the vehicle in a desired direction [7].

Other ways for user input can be electrooculography (EOG) techniques which reads the ocular position of the eye. The position of the eye can be used as a user input since it is possible to estimate where a user is looking. In [8] EOG was used to guide an autonomous wheelchair.

We have performed a user test where a hand held gyroscope unit controlled the semi-autonomous wheelchair of the MICA project. The gyro was based on an IMU equipped with three accelerometers and three rate gyros [9].

Different motions of the hand controlled the wheelchair in different ways, see Figure 1 for clarity. The MICA project has also studied ways to use a head mounted computer interface, MultiPos [10]. The MultiPos unit tracks the head motion of a user and moves the cursor on the computer screen accordingly. This means that a user can be immobilized but still be able to interact and drive the wheelchair with the interface if he can move his head.

2 The MICA Wheelchair-System Description

The MICA wheelchair is equipped with a PC in PC104 format running Linux. A fiber optic rate gyro (KVH E-core 1000) and two wheel encoders are used for pose tracking, Figure 3. A range scanning laser (SICK LMS200) is mounted tilted in the front to detect obstacles and find suitable paths, Figure 1.

2.1 Head Mounted User Interface

The head-mounted gyro is connected to the Laptop computer. The gyro interface tracks the head motions and moves the cursor on the screen. The "mouse"

Fig. 1. Mr Tapio Yliruikka, the wheelchair user, holds a steering unit that makes it possible from him to remote control the MICA wheelchair. The range scanning laser is visible on the wheelchair.

click can be done by inserting a bite button into the mouth, see Figure 3. The wheelchair moves as long as the button is pressed. Different commands can be selected on a user interface. The wheelchair stops when the bite button is released and a new command can be selected. Five different commands are described below. In semi-autonomous mode the wheelchair drive by information collected from the sensor units. The user can command the wheelchair to desired direction. The wheelchair automatically detects obstacles and finds a path around them. The semi-autonomous commands are:

Left. Drive left as soon as the there is a path available for the wheelchair to the left relative the wheelchair.
Right. Drive to the right as soon as there is a path to the right relative to the wheelchair.
Big. Follow the biggest passage.
Reverse. Reverse the wheelchair.
Turn 180^o**.** Turn the wheelchair 180^o degrees.

When the wheelchair is in motion the cursor will move on the laptop screen since the head mounted user interface is based on rate gyros. This is no problem since the wheelchair executes a command selected from the user interface as long as the mouse button is pressed. The bite button functions as a safety switch as the vehicle will stop as soon as it is released.

Fig. 2. MultiPos - A head mounted gyroscope used as a mouse interface to a computer. It is visible hanging on the laptop screen. With the gyro interface the user can operate the computer and select a destination for the wheelchair.

The user has click functionality through an installed Dweller program[11].

The head mounted computer interface can be used for autonomous movements of the wheelchair. The user marks a desired destination on a map, that describes the surrounding environment, on the screen. Using the laser technology, the wheelchair can detect the surroundings and calculate a collision free path. Figure 4 shows an example of a path in a typical corridor environment.

The on-board computer and its screen can also be used as a normal Internet connected PC running Windows XP.

3 User Test Using Hand-Held Unit

The testing was done at the Faculty of Art and Design at University of Lapland in Rovaniemi, see Figure 1.

The user sat in his own wheelchair and remotely controlled the MICA wheelchair over the Internet using hand motions only. The sensor unit was attached to the hand of the user and had a wireless serial link to the server computer [12]. The laptop collected user input from the server over the wireless Internet, processed the inputs and sent control commands back to the wheelchair.

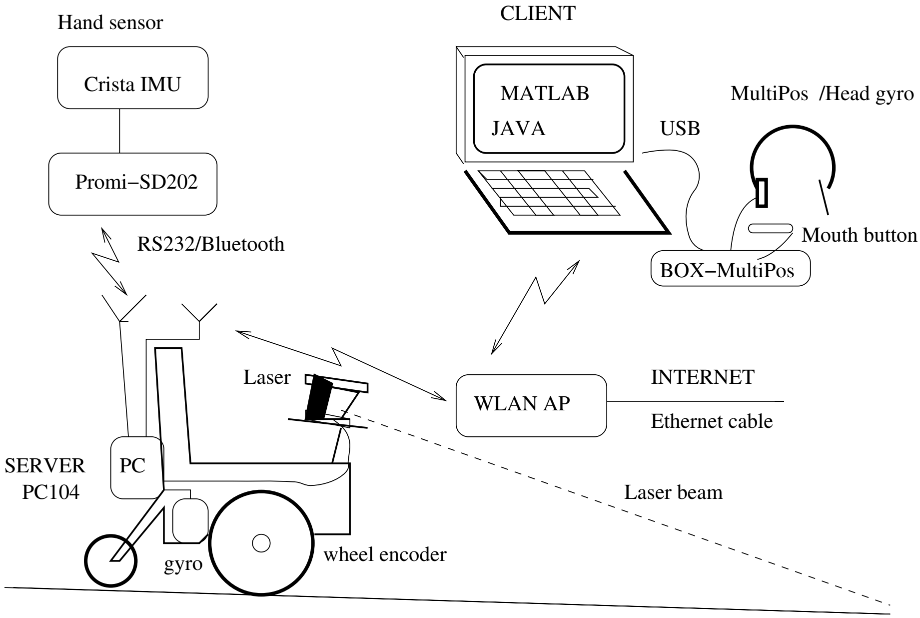


Fig. 3. System description. The wheelchair has a PC running Linux. It is connected to the Internet through a wireless network. The hand sensor (IMU) is connected to the wheelchair through a Bluetooth link. The head-mounted gyroscope is sampled by MultiPos box which is connected to the client computer. A button for the mouth is also connected to the box to give the user mouse button click functionality. A range scanning laser is mounted in the front of the vehicle.

3.1 User Opinion from User Test

We collected user comments and observed user behavior when he was testing the equipment. The user adopted easily to the hand held unit. Initially, the user had problems to understand how to use the hand control. According to the user the hardest part was to imagine how to move the control device as he was not sitting in the wheelchair.

The user of this test suffered from weak muscles and he had for example difficulties in lifting his arm. Twisting his hand, which was addressed to forward motion, was especially problematic for him. He said that the best would be to have a device or unit that was so small that it easily could fit on the top of his thumb.

4 Improvements and Ideas Based on User-Centered Design

Instead of focussing on sickness and weakness of elderly people, recent research brings up the issues like autonomy of elderly people and their relationship to their environments. There are also rights concerning elderly people and physically handicapped persons. These rights cover the possibility for persons to interact with the surrounding environment as well as having an environment that gives support to their needs in life. In practice, the MICA project produces information that can be used for developing user centered design. It is necessary, in a suitable way, to remove ageing based risks to make it possible for elderly people to continue their life at home. This means that a designer is encountered with a compromise between risks and the desire of elderly people to live at home.

Pressure on finding individual supporting environments for elderly people will grow. The MICA project answers to this pressure by gathering knowledge about users in order to produce a suitable design. Environmental design for elderly people demands a comprehensive design process. The designer can for example use collected knowledge from interviews to find out what kind of solutions are suitable.

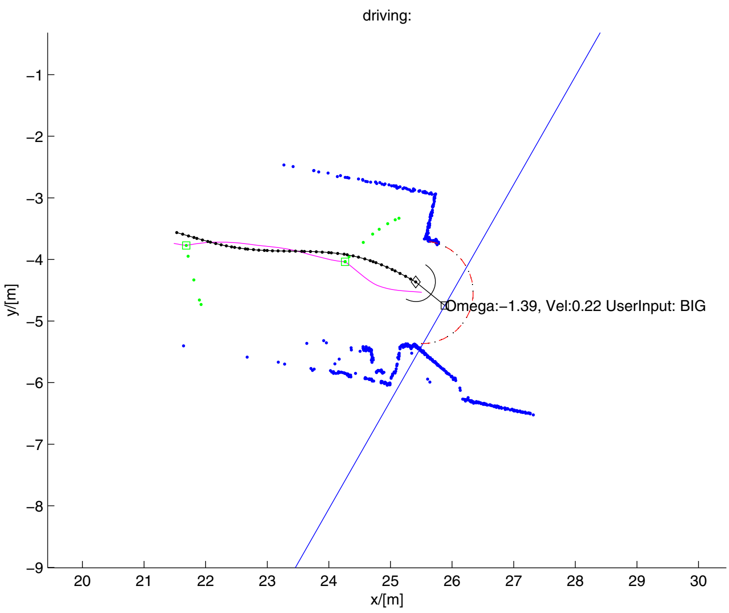


Fig. 4. Experiment with the wheelchair running in a corridor from left to right. The instantaneous pose of the wheelchair is marked with the small 180^o arc. Surrounding walls are visible on the left and right sides. Obstacle free region is marked with a dotted circle. Taken vehicle path is the dotted solid line. The long solid line marks the set boundary between the known and the unknown environment which is projected in front of the vehicle. The square on the line marks the steering direction. The 'BIG' text marks the input parameter from the user to the vehicle controller which says that the vehicle should move along the largest space available in front of it.

We use some basic ideas and assumptions. The first idea is that every person has a knowledge about himself, experiences from the own point of view, and that he influences the surrounding with them [13]. The second idea is that it is possible to find suitable methods for assistive improvements for people and their environments. The third idea is that, after getting knowledge about the user, it is possible to design a solution fitting the individual [14,15].

5 Results and Conclusions

The wheelchair technology can make it possible for a disabled user to navigate more easily in an environment.

The combination of technology, design, and social science form an important base for a practical development process. The MICA project stands as an example showing the importance of cooperation between complementary disciplines and the importance of sharing of knowledge between different faculties.

We have successfully used an IMU (inertial measurement unit), mounted on the hand of a user, to control a semi-autonomous wheelchair. Design is connected to social research and the contemporary research of ageing. This research, in turn, is looking for the answer to the question; What does the knowledge about a user mean to physical design? In the MICA project, knowledge about a disabled user have been used directly in technical design. It is possible to gather knowledge about a user, but the knowledge must be translated to an appropriate design language before it can be of practical use.When making design, especially for elderly people, memories are important to know. Hence, we consider technical design merged together with design based on, for example, personal wishes, favorite objects, and favorite places. The knowledge about a user is used as a resource in the design process.

Acknowledgement

The authors want to thank wheelchair user Mr Tapio Yliruikka in Rovaniemi for assistance in evaluating the steering method based on inertial sensors and Patrik Emilsson at Permobil for hardware support. The project was partly funded by the EU with Interreg IIIA Nord.

References

1. Luleå_University_of_Technology: The mica project home page. related data, movies and papers can be downloaded. http://mica.csee.ltu.se (2005)
2. Rönnbäck, S., Piekkari, J., Hyyppä, K., Haakapää, L., Kammunen, V., Koskinen, S.: Mica - mobile internet connected assistant. In: 1st International Conference in Lifestyle, Health and Technolgy, Luleå University of Technology (2005)
3. Harmo, P., Taipalus, T., Knuuttila, J., Vallet, J., Halme, A.: Needs and solutions - home automation and service robots for elderly and disabled. In: International Conference on Intelligent Robots and Systems, IROS'05. (2005)

4. Dautenhahn, K., Woods, S., Kaouri, C., Walters, M.L., Koay, K.L., Werry, I.: What is a robot companion - friend, assistant or bulter? In: International Conference on Intelligent Robots and Systems, IROS'05. (2005)
5. University_of_Lapland: Elvi - aiding the lives of the aging. University of Lapland, Rovaniemi, Finalnd, http://www.ulapland.fi/elvi (2004)
6. Bley, F., Rous, M., Canzler, U., Kraiss, K.: Supervised navigation and manipulation for impaired wheelchair users. In: IEEE International Conference on Systems, Man and Cybernetics. (2004)
7. Argyros, A., Georgiadis, P., P Trahanias, D.T.: Semi-autonomous navigation of a robotic wheelchair. Journal of Intelligent and Robotic Systems (2002)
8. Barea, R., Boquete, L., Mazo, M., Lopez, E., Bergasa, L.: Eog guidance of a wheelchair using neural networks. In: 15th International Conference on Pattern Recognition. (2000)
9. Cloud_Cap_Technology_Inc: Crista inertial measurement unit. http:// cloudcaptech.com/crista_imu.htm (2005)
10. Motion_Control_AB: Multipos. http://www.motioncontrol.se (2006)
11. Natural_Point: Smartnav. http://www.smartnav.com (2005)
12. INITIUM: Promi-sd product. http://www.initium.co.kr/english/promi-sd02.html (2006)
13. Varto, J.: Fenomenologisen tiedon kritiikki. The criticism of the phenomenological knowledge. Sufi researches. PhD thesis, Sufi tutkimuksia. Tampere (1992)
14. Jill, Q. In: Aging and the Life Course. An Introduction to Social Gerontology. McGraw-Hill College. Boston. (1999) 18
15. Cooper, R.A., Thorman, T., Cooper, R., Dvorznak, M.J., Fitzgerald, S.G., Ammer, W., Song-Feng, G., Boninger, M.L.: Driving characteristics of electrial-powered wheelchair users: How far, fast, and often do people drive? Phys Med Rehan **83** (2002) 250–255

Computer Control by Tracking Head Movements for the Disabled

HyungJun Kim and DaeHyun Ryu

Division of Information Technology
Hansei University, Korea
harry@hansei.ac.kr

Abstract. We present a system for the hands-free control of a computer, using a PC camera to track head movements, then translating these movements into cursor movements onto a computer screen. The proposed system is an alternative for people with spinal cord injuries and other special needs. It can be run on a standard Windows platform and makes use of a standard USB video camera to track head movements compared to other commercial hands-free products, which use special devices. Also, it can be applied to not only the disabled but also for any users operating common devices, such as ATMs, vending machines, or pay phones.

1 Introduction

Since many people use computers in everyday life, it is necessary to allow the disabled to communicate easier with computers. However, the disabled may have difficulties when controlling a PC using classical devices. They need special devices for controlling a mouse or typing on a keyboard. The control of a computer by tracking head movements is the subject of our project, and it has been studied and researched [1,2].

There are some commercial mouses on the market as a substitute for a normal mouse, such as CameraMouse, Quick Glance 2, SmartNAV, and QualiEye. CameraMouse is software for hands-free control of a computer using a video camera to track body movements [3]. Quick Glance 2 is a mouse replacement device that allows the user to place the mouse pointer anywhere on the screen simply by looking at the desired location [4]. A head tracking mouse, SmartNAV, allows complete control of a computer with natural head movements [5]. The device uses an infra-red (IR) camera to track tiny reflective dots on the part of your head. QualiEye gives simple hands-free control of the PC using a web cam for tracking eye movements [6]. In the MEMREC project, Fejtová, *et al.* have affixed a small camera to the head of a person and tracked the eye position in the coordinate system independent of head movement [7]. Mühlehner presented a research on a prototype of a laser-based, head-controlled input device [8]. Each of the presented system has its merits and defects. However, these systems need special devices and most of them are very expensive.

The proposed mouse using head movements is an alternative for people with spinal cord injuries, such as quadriplegia, muscular dystrophy, and other special needs. Anyone with controlled movement of his head can operate the proposed

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 709–715, 2006.

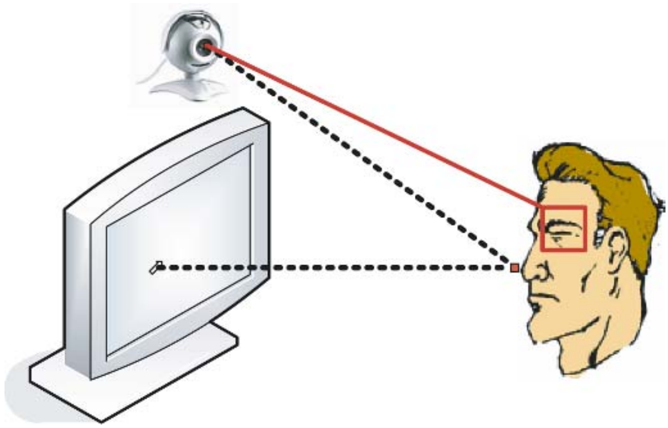


Fig. 1. Overview of the system

mouse. The system is very cheap and utilizes efficient MCI (Man-Computer Interface). A PC camera is connected to the PC, using standard communication interface equipped with corresponding software; therefore, the disabled can control a PC simply by head movements and an eye blink for clicking. Anyone with controlled movement of his head can operate the proposed mouse. Our research goal at the Center of IT Devices for the Disabled in Hansei University is the development of a simple and inexpensive system for the disabled that can be used for a PC mouse, as well as for a keyboard with some functions.

The paper is organized as follows: Section 2 explains the overview of the proposed system design. Section 3 describes the proposed method that predicts the motion region using neural network vector quantization and calculates motion vectors. Section 4 presents the usage of the system. Finally, Section 5 addresses the conclusions and future works.

2 System Design

We have designed a system that can be an alternative for classic mouses, using a PC camera to track head movements, then translating these movements into cursor movements onto a computer screen, using effective matching and tracking algorithms. The proposed system recognizes the position and horizontality of eyes and uses the information for clicking, double-clicking, and dragging processes. Therefore, the disabled can use the proposed system easily without any additional devices, except a standard PC camera. The software can be customized and adjusted for each user and was designed to accommodate varying degrees of control. Fig. 1 shows the overview of the system.

2.1 Template Matching

To track head movements and to translate these movements into cursor movements onto a computer screen, we use a small nose image as a template and compare the template with an input face-image sequence. One of the most fundamental means of object detection within an image field is by template matching, in which a replica of an object of interest is compared to all unknown objects in the image field. If the template match between an unknown object and the template is sufficiently close, the unknown object is labeled as the template object. However, a template match is rarely exact because of image noise, spatial and amplitude quantization effect, and a priori uncertainty as to the exact shape and structure of an object to be detected. An object is deemed to be matched wherever the difference is smaller than an established threshold value. Consequently, a common procedure is to produce a difference measure between the template and the image field. Also, we have to consider the processing speed to trace the template as fast as possible to avoid missing a user's movement.

For the template matching process, we assume the following features:

1. A nose is located at the center of a face.
2. There is abrupt gradient change at the nose area because of nostrils.
3. There is no sudden change during rotation and translation of a face.

The usual difference measure is the MAD (Mean Absolute Difference) or MSE (Mean Square Error), but these measures are just comparing gray level difference. Therefore, we applied NGC (Normalized Gray-level Correlation) for the matching process, that can compensate the gray level after subtracting the average gray value from the image itself as defined by Eq.1.

$$r = \frac{\alpha \cdot \beta}{|\alpha| \cdot |\beta|} = \frac{\sum_{i=0}^{M} \sum_{j=0}^{N} \alpha(i,j)\beta(i,j)}{\sqrt{\sum_{i=0}^{M} \sum_{j=0}^{N} \alpha(i,j)^2 \sum_{i=0}^{M} \sum_{j=0}^{N} \beta(i,j)^2}} \tag{1}$$

where N, M are sizes of the templates, $\alpha(i,j) = s(i,j) - \overline{s}$, $\beta(i,j) = t(i,j) - \overline{t}$, and s, t are the part of image and the template with its averages $\overline{s}, \overline{t}$, respectively.

We select a template at the center of the screen, which is 16×16 pixels. After registering the nose image as a template, the input image is scanned from top-left to bottom-right to find out the best matching position between the nose template image and the input face-image. To save searching process time, the search region is restricted within three times wider than the template image, which is therefore 48×48 pixels wide. Because the input image sequence is sampled fast enough, it can be guaranteed that the best matching position can be found within the search region.

2.2 Dual Template Matching

The template matching process is relatively stable if the head movement happens within a search region. However, if there is a sudden change of head movement, such as if a user turns around, the matching rate may drop rapidly. Therefore, we designed another small template, 10×10 pixels, as a supplement inside the original template image. The small template is for the tip of the nose and can be used in case the matching rate using the original template drops suddenly below 0.6. With a dual template matching process, using nose and tip-of-the-nose templates, we can track head movements more accurately even if there is a sudden head movement.

2.3 Clicking

We have utilized the gray level difference between two consecutive image frames to find out blink movements that can be translated to a click process. When a user blinks, there is a gray level difference around the eye areas, so we can find out whether or not a user blinks. The position of one eye may be determined, using the relative distance from the nose position as a reference, which is registered at the initial stage. Here the relative distance means that we can find local positions of the nose, eyes, and mouth in a face depends on human race.

We convert a color image that comes from a PC camera to a gray level image first, and then make it a binary image, using a pre-defined threshold value. Therefore, we can easily check the blink action by checking the amount of black portion around the eye area. When a user blinks, eyes close momentarily, so the iris area disappears briefly. We may check the black portion between two consecutive image frames and recognize a user's blink action. However, a user may blink unconsciously, so to discriminate an unconscious blink from an intentional blink, we treat double blinks within a short period of time as a one-click action. We have used the time interval difference to discriminate between normal blink movements and blink movements that should cause a click.

3 Motion Vector Estimation by Search Region Prediction

We propose a new method for estimating motion vectors in an image sequence. The proposed method predicts the search region by neural network vector quantization (VQ) and evaluates distortion for the predicted points. The simplest method of temporal prediction is to use the previous frame as the predictor for the current frame. Block Matching Algorithms (BMA) are utilized to estimate motion at a block of pixels (size of $(M \times N)$), for example, the location of the block in the present frame compared to that of the block in the previous frame. This block of pixels is compared with a corresponding block within a search region in the previous frame. The process of BMA divides an image into fixed size sub-images, and then finds one best match for the previous frame by maximizing cross correlation.

We define a function $D(\cdot)$ to locate the best match:

$$D(i,j) = \frac{1}{NM} \sum_{m=1}^{M} \sum_{n=1}^{N} G(U(m,n) - U_r(m+i, n+j)), \quad -p \le i,j \le p \quad (2)$$

where $G(\cdot)$ is a nonlinear function to evaluate error power, U is a block of size $M \times N$ in the current frame, U_r is a search area of size $(M + 2p) \times (N + 2p)$ in the previous frame, and p is the maximum displacement allowed. The value (i, j) is displacement which minimizes the $D(i, j)$ [9].

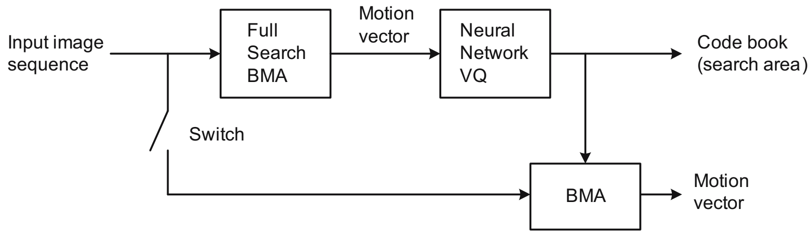


Fig. 2. Block diagram of the motion vector estimation system

The performance of motion vector detection can be increased because motion vectors usually have high spatiotemporal correlation. We propose a new motion vector estimation technique using this correlation. Fig. 2 shows the block diagram of the proposed motion vector estimation method using a neural networks vector quantizer. First, we find motion vectors using the full search method from the training images and then, train the neural network vector quantizer codebook using these motion vectors. Second, a motion vector can be estimated using the codebook as a motion prediction region. The codewords in the codebook represent the motion vectors for the input image sequences. Since the codebook is used as the search region for estimating the motion vectors, the search points and computation can be reduced compared with the full search BMA. The codebook is designed with the neural network vector quantizer utilizing the FSCL (Frequency Sensitive Competitive Learning) using the above motion vectors as the training input data [10].

4 Usage of the System

When a user runs the setup process, the user can put user's nose at the cross point on the screen and click the *Test* button as shown in Fig 3. It registers

Fig. 3. Example of the setup process

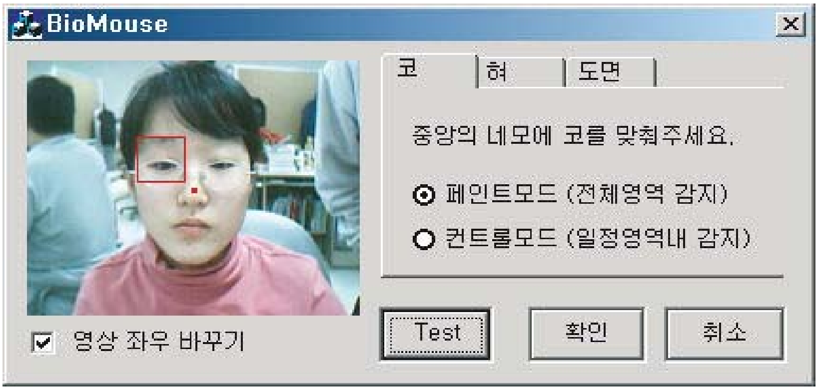


Fig. 4. Example of tracking head movements

a user's nose image as a template image. The system automatically detects a user's nose area and uses it for template matching as well as double template matching. When a user presses the *Start* button, the user can see a rectangular window around the right eye area with a red dot on the user's nose. This means the system is ready to trace his movement. When a user moves user's head, the rectangular window and the red dot follow user's movement as shown in Fig 4. If the user wants to single click or double click, the user blinks one eye twice or three times, respectively.

5 Conclusions and Future Works

We have designed a system that can be an alternative to a classic mouse, using a PC camera to track head movements, then translating these movements into cursor movements onto a computer screen. The proposed system was originally designed for the disabled to accommodate their usage of a PC. However, it can be applied to not only the disabled but also for any user because of its extensibility to PC games or ATM machines. The system cost is very low since there are no special devices required, except a standard USB PC camera. It is very simple, efficient, and easy to use. The proposed system recognizes the head position and interprets eye blinks as an ordinary mouse-click action. It also can be applied to an on-screen keyboard or a remote system to control on-off switches of electric equipment.

The points at issue of the proposed system are as follows: First, the system should register a user's nose image for the template image whenever a new user wants to use it, and it can be cumbersome work. We may apply an artificial intelligence algorithm to train the nose recognition process not to register a nose image as a reference. Second, when a user's face turns fast enough, the recognition rate drops. We may need special vectors for describing the position of the nose effectively. We have been working to resolve these problems. At the present it is possible to control a PC using the proposed system similarly to a standard PC mouse.

Acknowledgements

This work was supported by Hansei University.

References

1. Wu, T., Meng, L., Wang, H., Wu, W., Li, T.: Computer Access Assessment for Persons with Physical Disabilities: A Guide to Assistive Technology Interventions. in Computers Helping People with Special Needs LNCS 2398 Springer-Verlag Berlin Heidelberg New York (2002) 204–211
2. Betke, M., Gips, J., Fleming, P.: The Camera Mouse: Visual Tracking of Body Features to Provide Computer Access for People With Severe Disabilities. IEEE Trans. Neural Systems and Rehabilitation Engineering 10(1) (2002) 1–10
3. http://www.cameramouse.com
4. http://www.eyetechds.com/qglance2.htm
5. http://www.NaturalPoint.com
6. http://www.qualilife.com
7. Fejtová, M., Fejt, J., Lhotská., L.: Controlling a PC by Eye Movements: The MEMREC Project. in Computers Helping People with Special Needs LNCS 3118 Springer-Verlag Berlin Heidelberg New York (2004) 770–773
8. Mühlehnerer, M., Miesenberger, K.: HeadControl+: A Multi-modal Input Device. in Computers Helping People with Special Needs LNCS 3118 Springer-Verlag, Berlin Heidelberg New York (2004) 774–781
9. Iinuma, K., Koga, T., Niwa, K. and Iijima, Y.: A Motion-Compensated Interframe Codec. in Proc. Image Coding SPIE 594 (1985) 194–201
10. Al Sayeed, C., Ishteak Hossain, A.: Image Compression Using Frequency-Sensitive Competitive Neural Network. in Proceedings of the SPIE 5637 (2005) 611–618

Healthcare Service with Ubiquitous Sensor Networks for the Disabled and Elderly People

Yung Bok Kim[1] and Daeyoung Kim[2]

[1] Department of Computer Engineering,
Sejong University, 143-747, Seoul, Korea
yungbkim@sejong.ac.kr
[2] Information and Communications University (ICU),
119 Munji-Dong, Yuseong-Gu, 305-714, Daejon, Korea
kimd@icu.ac.kr

Abstract. An e-healthcare service with ubiquitous sensor network (USN) for the disabled and elderly people was studied, considering the current technology as well as forthcoming technology and service in the ubiquitous computing and networking environment. We introduce the USN for e-healthcare service for the disabled and elderly in smart environments. Beyond e-healthcare service, as a primitive application for ubiquitous healthcare service using mobile Internet, we studied the real-time health-monitoring service for the disabled and elderly people with an inexpensive and effective Web server. We considered the health-monitoring sensors in the wrist phone, as a future product for ubiquitous healthcare service. For quality of service (QoS), we studied an evaluation scheme for U-healthcare service for the disabled in smart environments, considering diversity of technologies and services.

1 Introduction

E-healthcare is a new paradigm in medical care on the basis of Internet and IT technologies, and it will improve our quality of life, especially in the Aged Society. Beyond e-healthcare service, the ubiquitous healthcare (U-healthcare) service will be more evolved and advanced medical service in the emerging ubiquitous society.

Let's discuss some researches related to smart environments and e-healthcare. For the visually impaired in educational environment, the use of the bar code and the RFID with reader-writer attached to a PDA was studied by Tatsumi et al. [1] as a promising step toward building an information ensured area, especially to get information from a bar code or from RFID tags attached to equipment or surroundings. 'Smart environments for all' was introduced to the special thematic session by Nussbaum [2], and he stressed that smart environments have a lot of potential to increase the quality as well as the efficiency of healthcare. Development of smart home technologies dedicated to people with disabilities provides a challenge in determining accurate requirements and needs in dynamic situations; and Feki et al. [3] introduced the integration of context awareness and multimodal functionalities in the smart environments. Information such as the availability of resources, user profiles, location, input

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 716 – 723, 2006.

controls and services can be used to improve the interaction between users and their environments.

As a promising technology that enables ubiquitous computing and leads IT (Information Technology) industries of next generation, the exciting new field of sensor networks is attracting so much attention and considered to be one of the hottest research topics these days. Today's sensor nodes (small devices with integrated sensor capabilities) can be efficiently powered with batteries or solar cells and run bi-directional communication to other devices forming ad-hoc networks which cooperate in data gathering and processing. In the emerging USN, sensor devices could be severely miniaturized, harvest energy from the environment and communicate with other networks and devices integrated in our homes or our cities as smart environments.

Recently, the new application of health monitoring and medical care based on USN is gaining popularity amongst researchers and gives good promises for future practical uses. We will introduce briefly this new field of work and the most active research in this area. We will also introduce a framework of USN for e-healthcare service in smart environments. We will describe some of the research in USN architecture illustrating how the work will contribute to the field of USN in general and e-healthcare for the disabled and elderly people in particular.

The five basic functions (5C) of e-healthcare service are content, community, clinical care, connectivity, and commerce. Ubiquitous healthcare can be included in a ubiquitous commerce in the future, on the basis of health information, tele-medicine, and tele-healthcare [4].

The following chapters are organized as follows. We will discuss the applications of USN in the field of health monitoring and e-healthcare in the evolving computing and networking environment, providing a brief introduction and a summary of the most important ongoing research. We will present USN architecture as a platform for e-healthcare service in smart environments, and discuss its functionalities and its applications. Beyond e-healthcare service, as a primitive application for ubiquitous healthcare service, we introduce the real-time health monitoring service using health-monitoring sensors and health information Web server in the mobile Internet before the proliferation of ubiquitous computing and networking environment. For quality of service (QoS), we introduce an evaluation scheme for U-healthcare service for the disabled in smart environments. Finally we conclude our study with consideration for further research.

2 e-Healthcare Applications

The need for health monitoring in e-healthcare service is gradually increasing and its application is becoming feasible with wireless sensor network technology. By adopting tiny wireless sensor network devices in the ubiquitous networking environment with some specific health monitoring system, regular patients and disabled people can be observed at smart environments providing e-healthcare service. For this purpose, wearable vital sign sensors can be attached to people's body, allowing continuous communication transferring people's sensed physical status. Health monitoring with ubiquitous sensor networks (USN) in the ubiquitous networking environment has some outstanding features comparing with traditional medical healthcare systems.

On the other hand, health monitoring based on USN in the ubiquitous networking environment provides a totally different healthcare system scenario. Sensor nodes are initially small and generally use battery rather than annoying power cables. In spite of their handy or tiny size, they can communicate with other sensor nodes, and consequently data from attached sensors is easily transferred to health monitoring systems in the mobile Internet. Monitoring of body status can be sensed and checked in a real-time manner without any notification or administrative managements. As an example, if the monitored patient becomes in a sudden emergency state, programmed sensor network actuators might automatically make emergency calls and the monitoring system would start sending video image along with sensed vital sign data to relevant medical doctor using wireless Internet. If required, sensor node networks can become large enough to cover whole hospital buildings or home areas with thousands of sensor nodes, in smart environments.

Sensor network's general features such as tiny sensor nodes, network construction and self-configuration make it possible to be used in medical care monitoring application. We present important requirements for healthcare monitoring sensor networks as follow.

- *Reliable data communication:* Most importantly, data communication should be reliable to send patient vital signs in real-time.
- *Wearable sensor nodes*: After combining health sensors into a hardware node, its size should be small enough to keep it convenient for wearable use.
- *Security*: Sensed data reveals personal information such as people's health information, habit and residential movements; for personal data reliable security mechanisms are required.
- *Sensor nodes mobility*: Along with user movements, sensor network topology will be dynamically changed. Thus sensor mobility needs to be considered suitably.

Before we introduce a USN as a platform for e-healthcare in next section, we discuss briefly some research projects related to healthcare. CodeBlue [5] is a famous health monitoring system using wireless sensor networks. CodeBlue integrates sensor nodes and other wireless devices into a disaster response setting to provide medical care monitoring. Another related research project is the wireless body area network for health monitoring developed by Jovanov et al. [6]. Their work aims to support ubiquitous and affordable healthcare. The main architecture consists of a three-tier ubiquitous monitoring system: wireless body area network, personal server and healthcare providing server.

3 A Platform for e-Healthcare Applications

A USN platform for e-healthcare service, an evolvable network of tiny sensors, constitutes the particular USN design for smart environments of healthcare service to address the challenges imposed by a dynamic future based on personal mobility. Contrary to other architectures which rely on static behaviors, ANTS (an evolvable network of tiny sensors by ICU in Korea) [7] is built with the idea of providing adaptability for dynamically changing environment, particularly coupling with the needs of the inherently dynamic healthcare monitoring systems with advanced features in medical care environments.

ANTS research team has been involved in several medical care research projects, such as the e-health project developed at the Information and Communications University in Korea. The system was designed to include wireless sensor networks at home environment to collect ECG data from the patient. In order to convey the data from wireless sensor networks at home to the Internet and to analyze and process the data, the project used wired and wireless networks including Internet.

The ANTS research team developed the sensor network field. The host attached to the ECG collector gets patient status information and forwards it to the sensor node, which sends the data to other sensor nodes across sensor network via RF. Similarly, another host receives the data from the sensor network and forwards it to the Web server using TCP/IP protocol suite in the Internet environment.

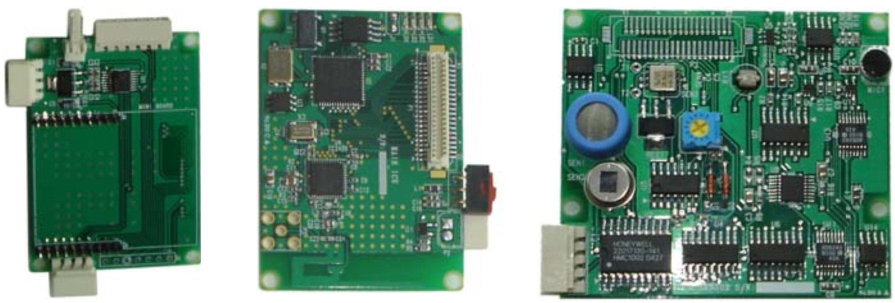


Fig. 1. (from left to right) Interface board, CPU board and Sensor board

ANTS architecture includes, besides software considerations, its own hardware support (as in Fig. 1) based on four different hardware designs suitable for different sensor network requirements. CPU board is responsible for monitoring and processing signals from devices such as ECG signal and heartbeat from Sensor board. For wireless communication to transmit personal health information, the design includes the 802.15.4 standard for radio communication, used in the Zigbee specification and offering promising reductions in power consumption. Sensors such as ECG, heartbeat, light, accelerometer and magnetometer are integrated into Sensor board. All these sensors can be operated individually or in collaboration with other sensors. If we want to use external health sensor or transmit collected health information to local computer system connected to Internet, we can use Interface board. It supports Serial/Parallel communication to communicate with external sensors or computers connected to Internet, i.e. wireless Internet or mobile Internet.

The operating system in USN is the core of the platform and coordinates the structural design to provide evolvable services. With low-power consumption, small code, small data size and evolvable architecture as design criteria, the operating system in USN platform was designed for wireless sensor network applications.

4 Ubiquitous Health-Monitoring Service

In the new and exciting field of Ubiquitous Sensor Networks (USN) we introduced, e-healthcare services for medical care applications appear as a promising research area with a wide range of possibilities. Beyond e-healthcare service, we introduce ubiquitous healthcare service. We introduce the health monitoring sensor network based on mobile Internet [8] as a primitive application for ubiquitous healthcare service.

4.1 Health-Monitoring Sensors

We can consider various types of health monitoring sensors for real application in the future, but here we considered the wrist phone as one of the future products to measure health data. Because oriental medical science is using information in the wrist to diagnose personal health, and we considered the real-time health-monitoring network based on mobile Internet. The accuracy and usability of the health-monitoring sensors will increase based on the evolution of the U-health technologies as well as on the evolution of MEMS (Micro Electronic Mechanical System) and USN technologies.

Glitho et al. [9] introduced a case study about the mobile agents and their use for information retrieval; the following health-monitoring sensors on wrist phones may be considered as mobile agents. Here instead of plain client-server, we need the optimized client server or the mobile agent approach for performance of our health-monitoring sensor network with health information Web server.

The performance analysis for the real-time health monitoring system may be different from the conventional analysis methodologies of non-real-time applications. We considered this type of evaluation, because the real-time health-monitoring network may be applicable to many new applications for the disabled and elderly in pervasive healthcare [10]. We studied the real-time health-monitoring application, getting frequently personal information as well as writing analyzed information as a response to the request by clients or health-monitoring agents in the mobile Internet environment.

For the real-time health-monitoring network using mobile Internet, the dominating factor and the standard deviation of that random variable should be bounded within the deterministic response time. To be deterministic for real-time application, the estimation time should be bounded within deterministic time, the interchange of data between the watch phone and the server should be automatic except the requested information by the user; therefore the Web server should be efficient and have high performance for the dedicated application if possible, and the exchanged data and analyzed information should be as simple as possible with a simplified and efficient format. If possible, the bandwidth requirement for wireless or mobile Internet should be immune to network traffic conditions; also that will be ideal in case of the degradation caused by the other rich multimedia contents sharing the same network and server. The spent time in the Web server may be considered to be immune to the network and server condition. This system is based on wired or mobile Internet, and the monitored health data from sensors in a watch phone can be registered at any time using mobile Internet.

4.2 Health Information Web Server

For the consistency of health-monitoring information and for a convenient user interface, we need a unified health-monitoring Web server for wired Internet and mobile Internet. We need to consider the health-information center accessibility to the doctor or nurse as well as to the disabled and elderly even with different formats of health information.

We used a single Web server as a health information Web server for cost-effectiveness and the simplicity of management. This method offers effectiveness and efficiency for the real-time health-monitoring network and utilization of resources, e.g. the bandwidth for communication and the size of disk storage of health information Web server for the disabled and elderly.

We assume that the wrist phone with health-monitoring sensors writes the health information regularly, which should be chosen carefully in further research, to the health information Web server through the mobile Internet, and processes the data and analyzes for the request of the results in real-time way. Depending upon the frequency of writing the health information, the workload of the health Web server may change, and the interval of regular writing may be considered as an arrival rate in the queuing performance analysis model for health-monitoring. The transmission packet unit for billing by mobile communication service provider is about 0.6 Cents (U.S.); the packet size is 512 Bytes, which is the minimum packet size for billing in Korea. Therefore, if possible, the health-monitoring data for the wrist phone to the server through mobile Internet should be below 512 Bytes, and this size is also bounded much below 1.5Kbytes that is one emulated WML deck for performance evaluation of the health-monitoring Web server discussed later. As a reference, one SMS (Short Message Service) data, i.e. 80 Bytes, costs around 2.5 Cents (U.S.) in Korea. Therefore, we can see that one WML packet is much cheaper than on SMS message in terms of cost for wireless communication.

The health Web server should have the capability to show the appropriate health contents, i.e. the HTML contents for wired Internet as well as the mobile contents for many different kinds of mobile devices, e.g. WML, mHTML, etc. For the unified service, there are several constraints, compared to the contents for the wired Internet. First of all, we should consider the various kinds of mobile devices as well as browsers for the mobile Internet, and each of those devices may have different capabilities in terms of images. The environment of the development and application is very different from the existing wired Internet, i.e. mainly based on MS Explorer. Therefore, we considered only text-based health-monitoring information from the wrist phone to the health-monitoring Web server and vice versa, to be immune to any type of Internet traffic load as well as to minimize the mobile communication cost for cost-effective health-monitoring services.

We suggest the minimum requirements as capabilities of health information Web server instead of superfluous capabilities because of the cost-effectiveness and low cost/performance ratio. To provide the minimum requirements for the wired Internet using a PC and for the mobile Internet using the wireless mobile devices, e.g. handheld devices and mobile phones, we can provide the simplest functionality with the less capable mobile Internet phones. We implemented the similar Web server of Information Network for the wired Internet with a PC as well as for the wireless Internet with mobile devices; however we considered the Web server as a health information Web server.

5 Evaluation Scheme for U-healthcare Service

We studied the important performance metric, delay, from the user's perspectives for the disabled. The preparation time for the disabled to get a service (e.g. medical device, mobile device, etc.) is *E* (environment time metric); the time spent by the disabled with medical device to do appropriate action for service is *D* (device time metric); the aggregate time to the medical healthcare server after the medical device for medical service is *S* (service time metric for communication and Web server); the time depending upon medical contents is *C* (contents time metric). If iteration of each variable is required, then subscript i,j,k,l,m can be used as E_i, D_j, S_k, C_l, P_m. Depending upon environment, device, service, and patient the iteration i,j,k,l,m may be different.

Among the above random variables, i.e. the performance metrics, (*E, D, S, C, P*) for the disabled, the most dominating factor, i.e. the random variable, may be different depending upon U-healthcare environment, device, service, contents, and patient. We can represent the statistical values (i.e. mean, deviation, etc.) of random variables as performance metrics for QoS (quality of service) of real-time medical service, and the quality of U-healthcare service can be compared with the performance metrics. Each performance metric can be used to enhance the QoS as follows. Environment time metric *E* will be shortened depending upon the proliferation of smart environment for U-healthcare. Device time metric *D,* i.e. handling speed of medical device, is one of important performance factors in any medical services for the disabled. Service time metric *S* can be shortened according to the communication facility and Web server with DB in e-hospital. Contents time metric *C* can be shortened considering accessibility in contents design for U-healthcare service. Patient time metric *P* may be different from person to person depending upon the disability of the disabled.

We can order the dominating factors in the overall performance at the user's perspective. The user interface design for the disabled with a medical device in a ubiquitous e-healthcare environment (i.e. *E* becomes smaller as the ubiquity increases) is important to decrease the device time metric *D* that is heavily related to the UI (user interface) convenience of medical device for the disabled and handicapped users.

With the performance metrics (*E, D, S, C, P*), we can estimate another metrics for QoS as follows: for example, total service time for a patient, T_m . We can use mean values of the random variables for simplicity of metrics as follows: $\overline{E}$, $\overline{D}$, $\overline{S}$, $\overline{C}$, $\overline{P}$ and $\overline{T} = \overline{E} + \overline{D} + \overline{S} + \overline{C}$.

6 Conclusions

We introduced the USN and a framework for e-healthcare service. A real-time health-monitoring network for the disabled and elderly people was also discussed for ubiquitous healthcare service on the basis of currently available service based on wired and mobile Internet. The evaluation of the health-monitoring Web server for the wired and mobile Internet can be applied to efficient investment for worldwide health-monitoring services for the disabled and elderly with the health-monitoring sensors in the wrist phone as a future product. For quality of service (QoS), we introduce an evaluation scheme for U-healthcare service for the disabled in smart environments. Ubiquitous healthcare service based on USN will be studied further for real-time monitoring and diagnosis for the disabled and elderly in the Aged Society.

References

1. Tatsumi, H., Murai, Y., Miyakawa, M., and Tokumasu, S.: Use of Bar Code and RFID for the Visually Impaired in Educational Environment. ICCHP 2004, LNCS 3118, pp. 583-588, 2004
2. Nussbaum, G.: Smart Environments for All Introduction to the Special Thematic Session. ICCHP 2004, LNCS 3118, pp. 889-890, 2004
3. Feki, M.A., Renouard, S., Abdulrazak, B., Chollet, G. and Mokhtari, M.: Coupling Context Awareness and Multimodality in Smart Homes Concept. ICCHP 2004, LNCS 3118, pp. 906-913, 2004
4. Ji, K.Y.: Healthcare in Ubiquitous Society. Jinhan M&B, pp. 300-305, 2005
5. Lorincz, K.: Sensor Networks for emergency response, IEEE pervasive computing, pp. 16-23, 2004
6. Jovanov, E., Milenkovic, A., Otto, C. and Groen, P.C.: A wireless body area network of intelligent motion sensors for computer assisted physical rehabilitation, Journal of Neuro Engineering and Rehabilitation 2005
7. Do, T., Kim, D., Sanchez, T., Kim, H., Hong, S., Pham, M., Lee, K. and Park, S.: An Evolvable Operating System for Wireless Sensor Networks, International Embedded and Hybrid Systems Conference, Singapore, May 10-13, 2005
8. Kim, Y.B. and Kim, J.Y.: Real-Time Health-Monitoring Network for Disabled and Elderly People. ICCHP 2004, LNCS 3118, pp. 377-384, 2004
9. Glitho, R.H., Olougouna, E. and Pierre, S.: Mobile Agents and Their Use for Information Retrieval: A Brief Overviews and an Elaborate Case Study. IEEE Network, January/February, pp. 34-41, 2002

Guiding Support for 'Way-Finding' in Unknown Buildings: Design and Evaluation

José M. Falcó[1], Roberto Casas[1,2], Álvaro Marco[1], and Jorge L. Falcó[1]

[1] TecnoDiscap Group.
University of Zaragoza, 50018 Zaragoza, Spain
{jfalco, amarco}@unizar.es
[2] Instrumentation, Sensors and Interfaces Group,
Technical University of Catalonia, Spain
rcasas@eel.upc.es

Abstract. We face the support to finding the way in an unknown building, such as hospital or city-hall for people with special needs (reduced vision, cognitive disabilities, physical disabilities), which forces not only to find the way, but to find the suitable one for specific capacities. Our goal is to enhance accessibility to public services and leisure in an autonomous and easier way. The system design uses building and personal context to find an adequate way to his/her destiny. Then it will guide the person through the way, monitoring the route followed, re-conducting when needed and giving the chance to change destination. This paper shows a brief description of the system and the two evaluation stages performed with their conclusions. A multidisciplinary team with engineers, gerontologists, psychologist and two special education schools are involved in specification, development and evaluation.

1 Introduction

Wayfinding is the process by which a person makes use of his/her spatial orientation to maintain a heading towards her/his target, regardless of the need to avoid or move around obstacles as walls or people. The successful coordination of actions within the perceived surroundings of a dynamic setting (e.g., a public building) usually requires the continuous wayfinding guidance provided from the environment with landmarks, signs and other references.

People with any kind of disability need to find their way in an unknown building, like hospital or city council. Commonly, signs indicate the target floor, but usually do not indicate how to get to it when having navigation restrictions (e.g. elevator for physical disabilities). If cognitive disability is present, it becomes an unapproachable challenge. Automatic guiding assistant could 'go along' with the person, providing guiding messages to users as they progress on their route. Computer "social assistants" can be used to simplify daily living and to enhance autonomy: GUIA, our guiding system, has this goal. Thus, we focus on affordable and effective computer technology.

GUIA calculates the best option depending on the building status and user abilities. If more than one viable path is found, it chooses the most secure. Many

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 724–731, 2006.

researches have offered their results in the area of training for navigation either in virtual reality [1,2,3] or in human reality [4]. Other experiments designed for evaluating several interfaces [4] and methods for navigation training [5] have been used as base of the field experiment we have designed.

As in the Goodman's work on landmarks [6], field studies are necessary because guidance systems performance depends upon the landmarks, the building's structure and other surrounding factors and cannot be tested in a laboratory setting. In this direction, Goodman also describes several quantitative and qualitative variables to measure the goal of the experiment [7]. Kjeldskov shows the advantages of field experiment evaluating the efficiency of those systems by conducting experiments in realistic use settings [8].

This paper also explores the efficiency of GUIA as an automatic guiding system which provides an aged and/or disabled person with independence and autonomy in unfamiliar buildings. First our indoors guiding system, GUIA, is described. Then the methodology and performance of evaluation experiments is specified. Finally, results and conclusions are shown, making proposals for further work.

2 GUIA's Description

GUIA is an indoors guiding system. It is designed to work with any kind of people and in any building, but our first focus —and most exigent too— is to guide people with cognitive, age, sensitive or physical disabilities.

GUIA needs to know the complete building map together with the special and common landmarks it has. We will call common landmarks those we can find at any public building, like signs and directories. Special landmarks will be those things in the building which are not thought as landmark itself but which may result appropriate for guiding, such as vending machines or doors in different colors or styles (glass, wood...).

The inputs to the guidance system will be the target place (or places) and the abilities of the individual to be guided.

GUIA uses an indoors position and orientation system developed and tested by the University of Zaragoza, which gives about 10 cm accuracy in absolute position and less than 10^{o} accuracy of user's head orientation: BLUPS (BLuetooth and Ultrasound Positioning System). However its design is independent of this system, so it could be used with any other emerging positioning technology.

With all this data GUIA calculates the path that the user can follow looking for feasibility and comfort. As the user goes on through the building the system sends guiding messages according to user's needs (so far voice and text messages are implemented) until s/he gets the target. During the guidance several special situations can happen, GUIA is prepared to cope with the following ones: (a) User loses her/his way – goes away from expected path: GUIA reorients the user from the current place so s/he reaches the right path or recalculates a new path from the current situation and guides the user through it. (b) User asks for an unexpected place to go before getting to the target: GUIA calculates a new path

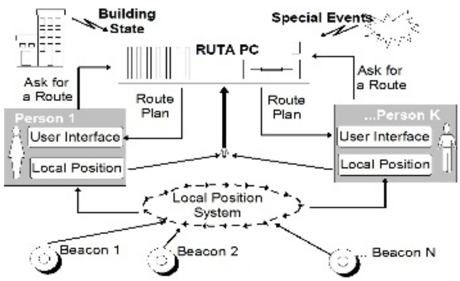


Fig. 1. GUIA's Block diagram

from the current place to the new target and from there to the main target. (c) User asks for repetition of last instruction: GUIA repeats last message. (d) An emergency situation happens in the building: GUIA guides the user to the nearest emergency exit. In Fig. 1 we show a block diagram of the system.

3 Experimental Method

We believe GUIA has a high potential to be used both as a guiding system and as a spatial orientation education tool. New experiments are being design together with special education teachers to provide a training tool in autonomous navigation. We want to find out the strong and weak features of GUIA. Besides, the users opinion will be important to evaluate the convenience of the guiding messages along the path, its temporal ratio (frequency) and how much understandable they were. This information is valuable for a further study on guiding messages.

Two experiments are designed and conducted: First emulates the system with a commonly used technique in sociological studies: an observer walking together with the user. The way was followed correctly, but we could not test the confidence the system generates in the user, for he/she feels accompanied along the path. Second experiment incorporates modifications that allow guided user to walk alone and receive recorded voice instructions, eliminating in such a way human contact, enhancing real autonomous feeling and allowing the test of the confidence the system may provide.

3.1 Scenario Description: Routes, Landmarks and Time of Day

Torres Quevedo Building is a three floor building with classrooms, offices and laboratories (Fig. 1, 2 and 3). We have designed a trip with three stages, to cover different considered difficulties. In Fig. 2, we show the following route and objectives.

First stage combines decision of route in open spaces and in confluences of corridors with 4 options, using stairs and elevator to change floor level twice, following long corridors. Goal is marked in Fig. 2 as "1", on the first floor (culture

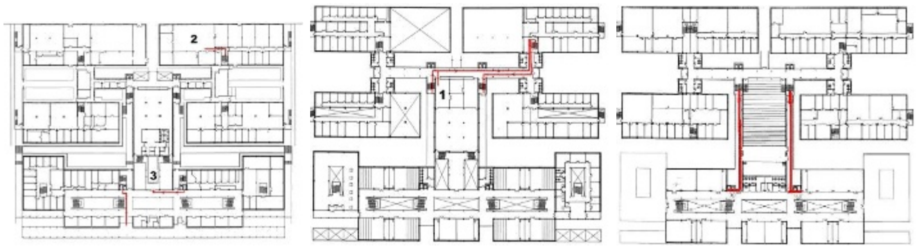


Fig. 2. Torres Quevedo Building. -0.5, ground and 0.5 floors with the routes marked.

room where many people students and from outside come to practice cultural and sports activities (dancing, meditation, aerobic). From the building front door the subjects come into an open space where they must find the elevator, get to the second floor, look for a sign pointing a door, follow a corridor and then, at the other side of the building, downstairs to first floor. Once again they come into an open space where they must find the door of the Culture room.

Second stage is chosen to walk through narrow and somehow labyrinth-like spaces, with many stairs that lead through three different ground levels ("1", "0,5" and "-0,5") with turns and short straight paths in between. Goal to reach is the precision mechanics workshop, visited by people from the building and "outsiders" as it is a research service. The subjects must follow a corridor, go across one door, downstairs to 0,5 level, turn, decide among three options, walk straight a bit and downstairs again to the hall on the -0,5 level where the workshop is.

The third stage leads to cafeteria, near the front door. Once again it is needed to find a sign which indicate 'EXIT', go upstairs, follow a corridor, find the elevator in a wide hall (open space) and then down to the ground floor. Finally in another wide hall they must find the cafeteria.

As example of changing dynamically the route, if any of the participants would want to go the WC there are several options all along the route.

We walked the entire route looking for elements which could be used as *"guiding landmarks"*. The Torres Quevedo building has its own signs pointing the different dependencies. We decided to use these signs as well as architectural references. As some of them are repeated in a short distance the guiding message talks about *'some EXIT signs'* and then *'we choose the first one'*. Some other elements were found too specific to be identified by common subjects, as an access control device, so they were refused as guiding landmark.

The experiments were performed in an active time of day, when lectures were in session and there were people moving around the building. Clearly making the experiment with an empty building would minimize external disturbances, but this way GUIA's efficiency is tested in a more real situation inside a public building.

3.2 Testing the Concept: People Emulating Technology – First Experimental Setup

First experiment was conducted to prove GUIA's efficiency as a personal automatic indoors navigation system which can help a person navigate through an unknown building. Participants that took part in first evaluation test were chosen to cover a wide spectrum of ages, from eight to seventy five, to test differences in perceptions and decisions taken by adults and children. Nevertheless none of them had disabilities and all were Spanish. Conclusions do neither test intercultural nor language differences. As Eckmann proposes, they are people who do not know at all the building [4]. Moreover, none of the subjects are involved in the development of GUIA.

Experimental Setup and System Emulation. Main objective of this trial is to test efficiency of GUIA in guiding, getting people to the goal minimizing spare paths and time out of the route. This building has neither location system nor communication network installed, so our experimental set up included the emulation of their performance with two persons:

– One experimenter —called "the system"— emulates the performance of GUIA's interface: he is inside the building with a computer running GUIA's algorithms that calculate the best path. He reproduces messages to the user when the system indicates and does not have visual contact with the subject GUIA is guiding. He speaks out the messages through a walkie-talkie to the guided subject. Two-way communication is enabled to allow specific requests considered in GUIA, as "repeat please", or "take me to the closest toilet".
– Another experimenter —called "the locator"— walks near the subject informing the system about the position and orientation of the user. As "the locator" cannot give the position of the subject with the same ratio and accuracy as BLUPS does (every 3 seconds, centimetres accuracy), all along the building there are numbered predefined positions. The messages the locator tells the system are like: 'leaving point 1, heading twelve o'clock', 'stop near point 3 heading six o'clock', 'going from point 7 to point 9' and so on. Twelve is north orientation, six is south. No other type of message is allowed to avoid transfer of emotional or additional information that GUIA won't have access to. Those messages are transmitted via another walkie-talkie pair that has only one way enabled.

Other experimenter —"the observer"— walks near the subject scoring the experiment and taking note of every significant person's spoken or nonverbal expression, timing, hesitating moments and whatever happens which may result relevant. This investigator does not mediate on the GUIA, he remains deaf-dumb all through the experiment.

This implementation of GUIA agrees to Goodman's rule *"Take care to give experimenter too many tasks"* [7]. "System", "locator" and "observer" do different simple tasks and none of them is overloaded. With a common set of maps of

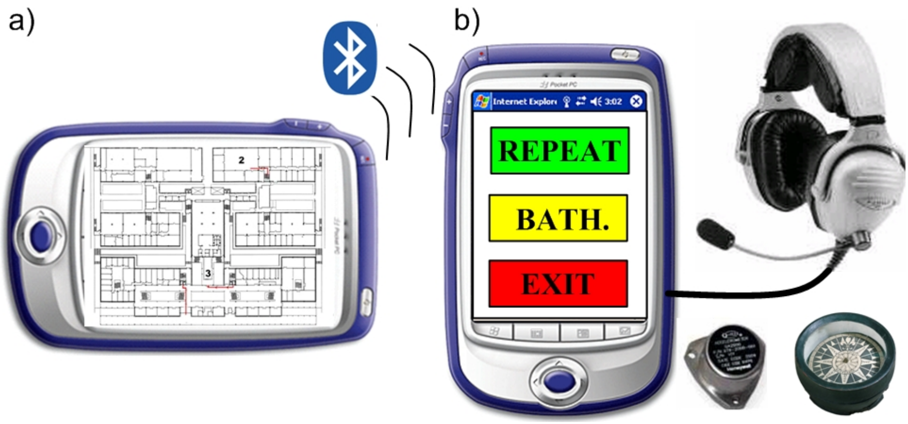


Fig. 3. a) PDA where the locator indicates and transmits the current user's position. b) user's PDA able to reproduce and record voice messages, process users' orders and save data from accelerometer and compass.

the building with special marks all along the trip, "the locator" reads the position of the subject, "the system" obtains the guiding message from the computer and "the observer" marks relevant events with place and time marks.

3.3 Testing Confidence: Only Technological Interfaces – Second Experiment

First experiment was conducted to prove GUIA's efficiency as a personal automatic indoors navigation system which can help a person navigate through an unknown building. Conclusions were considered valid as far as functional solution of the system, but not as far as confidence sensed by the user, for he/she had a person talking to him and this communication was felt as a personal relation with person "the system". Questionnaires showed they trusted the person could help them out in case it was needed. So a more technological tool was developed in which no human contact —even remote one— exists.

Experimental Setup and System Evaluation. In this experimental procedure participant's selection criteria and routes are the same as in former experimental setup. Observer was not present. The user brings a realistic interface that GUIA could have: a PDA where the guiding software runs (Fig. 3.b). Depending on his/her location and the received instructions (repeat message, new destinies —to the bathroom—, exit or take me out), it emits the appropriated voice messages. The user's interface has been designed the simplest possible: three colored buttons that input the former instructions to GIUA. The PDA also has a compass circuit and a 3 axis acceleration sensor connected to its serial port. All this allows recording the time, location, orientation and value of sensors in order to infer times of walking and stops, hesitations and differences in speed. The user is also asked to speak out loud specific comments or thoughts, which are recorded in the PDA for later analysis.

Location was communicated to the user's PDA via Bluetooth. A researcher was walking several meters behind with another PDA where he continuously set the user's location and orientation (Fig. 3.a). A wireless serial port among both PDAs was set in order to transmit all the information live. To make the experiment realistic, the person guided was not told about this person and he tried not to be related.

4 Results

According to Goodman's taxonomy [7], the wayfinding activity was scored using:

Timings: (a) Total time required to complete the entire path. In fact GUIA must guide the user to the target in a safe way, not as quick as possible. Thus, at the moment this is not a relevant measure, but it is easy to take and can be used in further studies. (b) Time taken to get the right way again if the user walks out of it.

Errors: (a) Number of times the subject leaves the correct path and comes back to it by her/himself. (b) Number of times the subject leaves the correct path and the system must correct the way. (c) Number of messages misunderstood. It is to say, messages that the subject mistakenly understands resulting in a different direction movement which message says. (d) Number of times reaching incorrect destination. We disagree with Parush and Berman because the natural end of using GUIA is reaching the correct target, so we must take note of the wrong ends [3]. If the user says 'Here I am' but s/he is not in the place the experimenters told her/him to reach, means the user has misunderstood some message or s/he is lost and GUIA has failed to get her/him where s/he wants to go.

Route taken: (a) Number of backtracking when the subject thought s/he was in the right way. (b) Alternative routes the system has offered when the subject has left the original route.

Other variables of interest: (a) Number of times the subject hesitates looking for the landmark the system is pointing. (b) Number of times the subject needs extra help and demands another message. (c) User comments: subject is told to think aloud while the experiment goes on and after the experiment is asked to complete a questionnaire. (d) Various experimenters' observations.

Major result is all the subjects reached the three targets. Latency in instruction arrival was observed in first experiment, what forced users to decide on their own or wait. Some re-conductions were debt to this factor. Only due to lateralization problems was route non-following detected, but it was quickly re-conducted by GUIA. To increase non-following, wrong messages were tested and once the users were out of the way, GUIA took them back easily to the planned route. In one point the wrong message was impossible (walking into a wall): The subjects stopped a bit surprised and waited for the right message. Most requests for repetition of guiding message were due to noise in communications.

In both experiments, and despite of intended wrong messages and misunderstandings, none of the subject felt lost at any moment and they expressed they relied on GUIA to get the planned targets. They also expressed the landmarks GUIA used were enough and appropriate. The talking PDA with three colored buttons to capture user preferences proved to be a valid interface even for the elderly people.

5 Conclusions

We have shown the description of our guiding system for people with disabilities including a location system and the special context information of landmarks to reference routes to the user. Also, how it has been emulated for testing in two experimental setups to gradually acquire data of efficiency and specific changes needed in design. First experiment gave conclusion of reliability in guiding, but lack of data in user confidence in the system, crucial for usability. Second experimental setup allowed detecting confidence status and orient changes if needed, before definitive implementation. Although elderly people took part in the experiments, confidence is to be proved also with people with disabilities. Then we may have a useful tool for special schools spatial education training and support to dependencies finding in large or complex buildings, which we trust will improve quality of life for people with disabilities and normalized people.

References

1. Bliss, J.P., Tidwell, P.D.: Effectiveness of virtual reality for administering spatial navigation training to police officers. Proceedings of the Human Factors and Ergonomics Society (1995).
2. Bliss, J.P., Tidwell, P.D., and Guest, M.A.: The Effectiveness of Virtual Reality for Administering Spatial Navigation Training to Firefighters. PRESENCE, Vol. 6. (1997) 73–86.
3. Parush, A., Berman, D.: Navigation and orientation in 3D user interfaces: the impact of navigation aids and landmarks. International Journal of Human-Computer Studies, Vol. 61 (2004) 375–395.
4. Eckmann, M., Yu L., Terrance E. B., Kessler, G. D.: Interfaces for navigation and familiarity training. Proceedings of International Workshop on Presence, (2001).
5. Darken, R.P., Banker, W.P.: Navigating in Natural Environments: A Virtual Environment Training Transfer Study. Proceedings of VRAIS. (1998) 12–19.
6. Goodman, J., Gray, P., Khammampad, K., Brewster, S.: Using landmarks to support older people navigation. Mobile HCI. (2004) 38–48.
7. Goodman, J., Brewster S., Gray,P.: Using Field Experiments to Evaluate Mobile Guides. Proceedings of HCI in Mobile Guides. 2004.
8. Kjeldskov, J. and Graham, C.: A review of Mobile HCI Research Methods. Proceedings of Mobile HCI 2003, Springer-Verlag, Vol. 2795. (2003) 317–335.

Mobile Computing in Medicine: Designing Mobile Questionnaires for Elderly and Partially Sighted People

Andreas Holzinger[1], Peter Sammer[2], and Rainer Hofmann-Wellenhof[3]

[1] Institute for Medical Informatics, Statistics & Documentation (IMI),
andreas.holzinger@meduni-graz.at
[2] Institute for Hypermedia Systems (HMS),
Joanneum Research, Inffeldgasse 16c, A-8010 Graz, Austria
peter.sammer@joanneum.at
[3] Department of Dermatology,
Medical University of Graz, Auenbruggerplatz 2, A-8036 Graz, Austria
rainer.hofmann@meduni-graz.at

Abstract. At the clinical department of Dermatology at the Medical University Hospital in Graz, approximately 30 outpatients consult the pigmented lesion clinic each day. During the visit, the patients are asked to complete a questionnaire, which is necessary, both for the clinical information system and for a scientific database for research in skin cancer. However, motorically and visually handicapped people usually have problems in completing paper based questionnaires. Consequently, a system was built, using a mobile touch computer with a specially designed interface, in order to assist these people and to allow full mobility within the clinical department, as well as the possibility of completing questionnaires, for example: during a cancer survey even in the open-air swimming resort. The system was developed by applying a User Centered Design including four levels: paper mock-up studies, low-fi prototypes, hi-fi prototypes and the system in real life. Scientifically this work provided insights into the technical possibilities, Human-Computer Interaction and Usability Engineering, user acceptance in the clinical field and the possible optimization potential of clinical workflows.

Keywords: Human-Computer Interaction & Usability Engineering, Mobile Computing, Information Interfaces, Input Devices and Strategies (mobile touchscreen), Screen design, User-centered Design & Development (UCD).

"Human-Computer Interfaces should not only support more effective and efficient user interaction, but also address the individual end-user requirements and expectations in the variety of contexts of use to be encountered". [1]

1 Introduction

Mobile computers generally present a number of challenges in Human-Computer Interaction (HCI) including the interplay between the appropriate user interface

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 732 – 739, 2006.

design, the device and the social context of the device's use [2,3]. The usefulness of mobile applications in Health Care is commonly accepted [4,5].

To achieve maximum benefits by making both useful and good usable applications, particularly in light of new devices, it is strongly recommended to apply a User-Centered Design (UCD) approach [6]. Some key principles of UCD methods include understanding the users and analyzing their tasks; setting measurable goals and involving the end-users from the very beginning. Based on the experiences within this project and on previous work [7,8,9] we found again that UCD is of particular importance to realize usable and useful applications, especially for mobile devices in such a difficult environment as an outpatient clinic. We cannot too often emphasize, that simple, cheap and easy-to-use solutions can be a step closer to the information society for all, where people are assisted by Information Technology [10].

2 MoCoMed-Graz: Technological Background

As part project of the project Melanoma Pre-care/Prevention Documentation (in German: Melanomvorsorgedokumentation), which is an important step toward fighting skin cancer, the project MoCoMed dealt with the design, development and implementation of a fully functional mobile solution to assist patient data surveys. The initial problem was, that the paper based questionnaires had several disadvantages; including the necessity of retyping them manually into the database, most of all, they were awkward to fill out by elderly and partially sighted patients, or for example by patients with tremor.

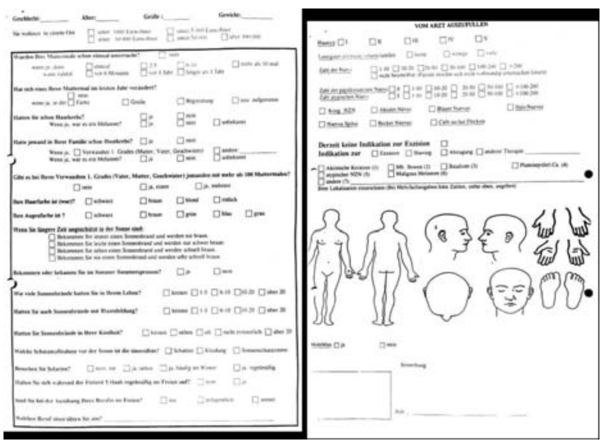


Fig. 1. The original paper questionnaire before this study took place

The initial idea of the use of mobile computers was to ensure that the data acquisition within the clinical department runs smoothly and also that the cancer researcher is allowed to collect data away from the clinic, for example during a survey study in an outdoor swimming pool.

We used an XML interface as the technical protocol, because the data collected by MoCoMed are directly transferred into the countrywide Hospital Information System (which is called MEDOCS and is a customized SAP-product [11]).

The workflow: The patient reports to the central administration desk of the outpatient clinic of the dermatology department. There, they are registered via the MEDOCS administration program into the pigmented lesion outpatient clinic.

At the clinical workplace, an overview of the waiting patients, who have been registered already in the system, but not yet released by a medical doctor, can be seen. In the corresponding column on the clinical workplace, there is an indication of whether or not they have already filled out a questionnaire. This is indicated by means of text and/or a symbol and makes the following differences visible: A questionnaire was filled out on the current day; A questionnaire has been made available to the patient, but not yet completed; A questionnaire, completed by the patient during a previous treatment is available (not from the current day); No questionnaire is available (column is empty). Now the medical doctor or the nursing staff of the clinic can decide whether this patient is to fill out a questionnaire and/or which questionnaire to provide to the patient. After the decision to ask the patient to fill out a questionnaire, an empty questionnaire is created in MEDOCS, by pushing a button. The questionnaire in MEDOCS is registered with a definite user and a unique identification code, so that it is clearly evident that it corresponds to a version from the patient and not the medical doctor. Using an XML communication, the patient identification number (PID), the unique number of the document (document number at the top of the questionnaire) and any further data (e.g. name, date of birth) considered necessary by MoCoMed, are transmitted.

At the terminal, the patient is equipped with a touch based Tablet PC and a code, with which he/she can login to MoCoMed and complete the questionnaire following a touch based application. The authentication at MoCoMed is necessary for data security reasons, so that no patient can access other data and patients avoid mistakes or errors. After the questions have been answered and the questionnaire is completed, MoCoMed transfers the questionnaire into MEDOCS. The corresponding column in MEDOCS now shows the status "questionnaire was filled out on the current day". This process must take place with the minimum possible delay, in order that the workflow of the outpatient clinic can continue undisturbed. As soon as the patient completed the questionnaire the XML file containing the answered is stored on the server, subsequently transferred to MEDOCS by using a remote function call (RFC). The XML document containing all answers of a patient includes of course the unique identification of each question. Further technical background of MoCoMed is not described in detail here, since we concentrate in this paper on the user-centered development including the scientific experiments and their findings.

3 MoCoMed-Graz: User Centered Design and Development

The first step in this project was to determine the project requirements and the clinical context. This is highly important in order to identify the end users. It is necessary to differentiate between the primary end-users, the secondary end-users and the stakeholders. However, the stakeholders influence, or are influenced by, the system but are not the actual users. A precise specification of end users is unavoidable, which includes the typical end user characteristics, e.g. age range, computer literacy and physical limitations (disabilities). Within a clinical development it is necessary to adapt the usability of the system to the lowest common denominator.

Technical Environment. At first we proposed the use of touch based Tablet PCs; however, the fear of problems, including petty larceny, destruction, misuse or misunderstanding, emerged immediately, particularly with regard to the lack of keyboard. Stakeholders proposed to solve this by installing a kiosk touch-station [12]; however, while this would solve some problems, it would totally disable our goal of total mobility. The solution was a trade-off, which later proved to be an absolute optimum: We decided to build a client-server system with a thin client solution at the front-end using a highly mobile device (Skeye-Webpanel). In order to reduce the danger of theft, the mobile device was mounted on a specially designed adjustable wheel table, whenever the medical staff leaves the patient to complete the questionnaire in private.

Physical Surrounding. The aim was to capture as much as possible information about the future workplace and physical conditions. Actually, the work atmosphere within an outpatient clinic is difficult, hectic and chaotic. For example, the noise level made several ideas of providing audio feedback inappropriate. Also, both low and high levels of lighting have an impact on end-users (office versus outdoor swimming pool, where sunlight is always a problem and causes glare on the screen). However, we also considered room and furniture because the characteristics of the installation place must be studied in order to operate the system safely and comfortably. We also carefully studied user posture; in our case it is possible to use the mobile device within a total mobile setting or on the adjustable wheel table (e.g. sitting versus standing and looking down at a display).

Context. The social and organizational context is most often neglected but is essential for the success of any system. Some factors within our project included general structure (e.g. hours of work, group working, job function, working practices, assistance, interruptions, management structure, communications structure, IT policy, organizational aims), and also attitudes towards the system, as well as work characteristics (e.g. job flexibility, performance monitoring and feedback, discretion, valued skills).

It is important to consider the staff and management structures in which the (proposed) system will operate.

Especially, the role of the end users must be considered, especially with respect to new procedures associated with the system, thereby maintaining level of status and activity satisfaction. When a new system is implemented, users often require support in learning how to use it and in solving problems. The possibility of offering support should be calculated and suitable support mechanisms should become part of the user requirements specification.

In our case privacy was also a key issue, and considerations that end users want to feel safe and secure in performing their tasks. If the system does not give the impression of safety (avoidance of loss of privacy) and security (assurance that only the medical staff is accessing the information provided) users will not perform well.

Level 1: Lo-Fi Prototyping: Paper Mock-Up. The term paper mock-up means "to prototype the screen designs and dialogue elements on paper" [13,14,9]. It again proved itself an easy and efficient method. With standard office supplies, each interface element (e.g. dialog boxes, menus, error messages, sliders (see figure 3)) has been sketched. This led to an easy creation of alternatives since it encouraged more suggestions due to the ease of alteration. We performed studies at this level with N=10 different people (see e.g. figure 2), the experiments were continued until no further findings were to be gained.

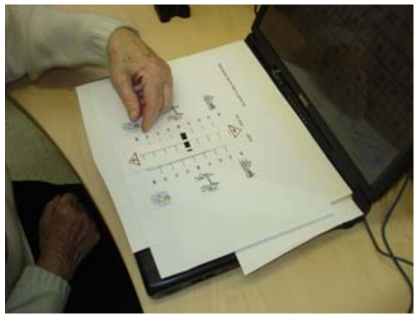


Fig. 2./3. One of our elderly patients is operating the paper mock-up (left), various input possibilities have been tested on paper (right)

Level 2: Hi-Fi Prototyping. In their experiments, Virzi, Sokolov and Karis (1996), [15] discovered that low-fidelity and high-fidelity prototypes have substantially the same set of usability problems. We can only partly agree: during the paper mock-up studies, we discovered more interaction difficulties than during the hi-fi level, where we found more difficulties and problems concerning the content, i.e. the wording and understanding of the questions. However, we can recommend low-fidelity prototypes unless performance measures are required.

Level 3: Real-Life. It is essential to test the final version in full operation within real-life.

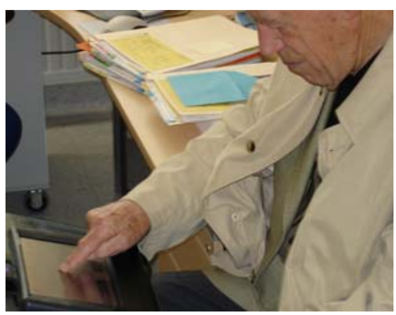


Fig. 4./5. An elderly patient is operating a Hi-Fi Prototype with full functionality; (left), Finally the SkeyePad in real-life operation (right)

4 Lessons Learned: Designing for Elderly and Partially Sighted

Although developing applications for mobile computers is considered to be different than for desktop computers [16,17,18,19], many general guidelines and experiences from desktop interfaces, especially experiences from touch-based interfaces [7,12,9] as well as and general usability engineering methods [20] are applicable.

The intensive study of end users by the application of paper mock-ups resulted in a great advantage and clear benefit. Some advantages were that the first sketches allowed immediate usability feedback. In the beginning of our project we were able to concentrate on abstract interface concepts and not on technological details. It further proved to be inexpensive to produce with a maximum feedback for minimum effort. However, as usual, some disadvantages also appeared: It was relatively difficult to simulate interface behavior and, in combination with the applied thinking-aloud and video-analysis, it needed far more time than was theoretically predicted e.g. according to Beaudouin & Mackay (2003) [21], because they failed to include the preparation time and the post-editing and post-processing time as well.

The high-fidelity prototyping had the advantage, that the end users could be studied in a more realistic setting (users could work with it directly), however, this scenario still does not adequately represent the final system. Concerning the design of the content of the questionnaire, we found that until the final experiments there were iterative improvements possible, including words, phrases and familiar - in the sense of intuition - concepts. It was interesting to observe that the end users chose system functions by mistake. Consequently, they also needed a clearly marked emergency exit button to leave the unwanted screen. We also made everything consistent avoiding different interface elements including words and actions. However, we followed an aesthetic and minimalist design: all dialogues contained no irrelevant information whatsoever.

According to the insights into end users behavior, we also built a special help function (figure 5), wherein, if there is no user input for a certain amount of time, a graphical hint (red arrow) encored the end users to touch the next button. Initially, we thought about audio feedback but the noise within the clinical environment made this inadvisable. The method Thinking Aloud (TA) revealed – with a fairly small number of end users (N=10) – why end users preferred certain interactions, consequently we could optimize both interaction and content. Especially early clues definitely helped to anticipate and trace the source of problems in the early stage of designs in order to avoid later misconceptions and confusion. Disadvantages included: nearly all the people investigated perceived this method as strenuous, it took a lot of time and preparation, and still more than 50% refused to voice valuable information out loud. With this method, there is always the danger that previously computer illiterate end users generally feel inhibited, which consequently slows down the thought processes, thus increasing mindfulness (which is normally good, but creates a bias because it may prevent errors which otherwise would have occurred in actual use). Generally these experiments were time consuming since it was necessary to prepare the end user with a careful briefing. Here, it is interesting to note that some elderly end users refused even to take part in the experiments when they heard the word “computer”; however, if we emphasized that we only wanted to test a newly developed questionnaire, the people were more likely to take part. This is possibly due to the fact that elderly people have had less exposure to computers and therefore have a total misconception of the capability of computers. Generally users interact differently with mobile devices than with PCs (see for example [22,23,24]). This does not quite apply to our case, since we used in our application a device with a resolution of 800 x 600 pixels. However, it is absolutely necessary to reduce the text input to a minimum, it is much easier to select values from a list rather than enter text in an input field.

Final example: One elderly lady within an experiment first were very frightened and said that she does not like computers at all – after she touched-through the application, she asked carefully: "This was a computer?", as we replied yes, she said "… but this was really funny". This proves that the work we do, is important in order to support an information society for all … however there is still much work to do!

Acknowledgements

We are most grateful for the technical support of Siegbert Kaiser, Herbert Kogler, Rene Malek, Peter Krasser, Andre Perchthaler, Daniel Zenz and Udo Pichl.

References

1. Stephanidis, C., ed. User Interfaces for All: Concepts, Methods and Tools. 2001, Lawrence Erlbaum: Mahway (NJ).
2. Myers, B., et al., Strategic directions in human-computer interaction. ACM Computing Surveys (CSUR), 1996. 28(4): p. 794-809.
3. Holzinger, A. and M. Errath, Designing Web-Applications for Mobile Computers: Experiences with Applications to Medicine, in User-Centered Interaction Paradigms for Universal Access in the Information Society. LNCS Vol. 3196, C. Stephanidis and C. Stary, Editors. 2004, Springer: Berlin, Heidelberg, New York. p. 262-267.
4. Bludau, H.-B. and A. Koop, Mobile Computing in Medicine, Second Conference on Mobile Computing in Medicine. Lecture Notes of Informatics. 2002, Heidelberg: GI.
5. Hameed, K., The application of mobile computing and technology to health care services. Telematics and Informatics, 2003. 20(2): p. 99-106.
6. Norman, D.A. and S. Draper, User Centered System Design. 1986, Hillsdale (NY): Erlbaum.
7. Holzinger, A., User-Centered Interface Design for disabled and elderly people: First experiences with designing a patient communication system (PACOSY), in Lecture Notes in Computer Science. Vol 2398, K. Miesenberger, J. Klaus, and W. Zagler, Editors. 2002, Springer: Berlin et al. p. 34-41.
8. Holzinger, A., Experiences with User Centered Development (UCD) for the Front End of the Virtual Medical Campus Graz, in Human-Computer Interaction, Theory and Practice, J.A. Jacko and C. Stephanidis, Editors. 2003, Lawrence Erlbaum: Mahwah (NJ). p. 123-127.
9. Holzinger, A., Application of Rapid Prototyping to the User Interface Development for a Virtual Medical Campus. IEEE Software, 2004. 21(1): p. 92-99.
10. Stephanidis, C. and A. Savidis, Universal Access in the Information Society: Methods, Tools and Interaction Technologies. Universal Access in the Information Society, 2001. 1(1): p. 40-55.
11. Gell, G., et al., HIS purchase projects in public hospitals of Styria, Austria. International Journal of Medical Informatics, 2000. 58-59: p. 147-155.

12. Holzinger, A., Finger Instead of Mouse: Touch Screens as a means of enhancing Universal Access, in Universal Access, Theoretical Perspectives, Practice, and Experience. Lecture Notes in Computer Science Vol. 2615, N. Carbonell and C. Stephanidis, Editors. 2002, Springer: Berlin, Heidelberg, New York. p. 387-397.
13. Wilson, J. and D. Rosenberg, Rapid Prototyping for User Interface Design, in Human-Computer Interaction, M. Helander, Editor. 1988, North Holland: Amsterdam. p. 859-875.
14. Rettig, M., Prototyping for tiny fingers. Communications of the ACM, 1994. 37(4): p. 21-27.
15. Virzi, R.A., J.L. Sokolov, and D. Karis, Usability problem identification using both low- and high-fidelity prototypes. CHI Conference on Human Factors in Computing Systems, 1996: p. 236-243.
16. Weiss, S., Handheld Usability. 2002, New York: John Wiley and Sons. 292.
17. Karampelas, P., D. Akoumianakis, and C. Stephanidis, eds. User Interface Design for PDAs: Lessons and Experiences with the WARD-IN-HAND prototype. User Interfaces for All. Lecture Notes of Computer Science Vol. 2615, ed. N. Carbonell and C. Stephanidis. 2003, Springer: Berlin, Heidelberg, New York. 474-485.
18. Buchanan, G., et al. Improving mobile internet usability. in International World Wide Web Conference. 2001. Hong Kong.
19. Holzinger, A., Multimedia Basics, Volume 3: Design. Developmental Fundamentals of multimedial Information Systems. 2002, New Delhi: Laxmi Publications.
20. Holzinger, A., Usability Engineering for Software Developers. Communications of the ACM, 2005. 48(1): p. 71-74.
21. Beaudouin-Lafon, M. and W. Mackay, Prototyping Tools and Techniques, in The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications, J.A. Jacko & A. Sears, Eds. 2003, Erlbaum: Mahwah (NJ).
22. Chittaro, L.E. Human-Computer Interaction with Mobile Devices and Services. in 5th International Symposium, Mobile HCI 2003, Lecture Notes in Computer Science, Vol. 2795. 2003. Udine, Italy.
23. Jessup, L.M. and D. Robey, Issues and challenges in mobile computing: The relevance of social issues in ubiquitous computing environments. 2002. 45(12): p. 88-91.
24. Brewster, S., Overcoming the Lack of Screen Space on Mobile Computers. Personal and Ubiquitous Computing, 2002. 6(3): p. 188-205.

NeurOSS — Open Source Software for Neuropsychological Rehabilitation

Jyrki Rissanen

Media Lab
University of Art and Design Helsinki UIAH,
Hämeentie 135 C, FIN–00560 Helsinki, Finland
Jyrki.Rissanen@uiah.fi

Abstract. In recent years hundreds of successful community-driven open source software projects have incarnated. However, it is quite hard to find similar success stories in the field of neuropsychological rehabilitation. This paper describes the core ideas of the NeurOSS project. The project aims at building an open source software platform for developing tools for neuropsychological rehabilitation, and gathering up a community of people from all over the world to contribute to shared collection of open source plug-in components extending and utilizing the core services of the platform.

1 Introduction and Aims of the Project

It is justified to say that there is hardly any multi-purpose open source software platforms available in the field of neuropsychological rehabilitation. Most of the products in the market have been made by U.S. companies, and those companies have localized their products only for the biggest market areas. Almost all the available products are like black boxes without any possibility for localizing software for different languages, or to modifying and extending the functionality of it. It is worth mentioning that most of such products are simply too expensive for the health care organizations in 3rd world countries [1].

The main motivation for initiating the NeurOSS project has been to address above problems by creating a carefully documented open source platform offering a well-known, widely used plug-in architecture for developers to extend the functionality of the platform for needs of the neuropsychological rehabilitation. The first version of the platform will be available for public by the end of 2006. At the same time, a few freely modifiable plug-ins related to dyslexia will be published, demonstrating the possibilities of the NeurOSS platform in real neuropsychological rehabilitation tasks.

The NeurOSS platform with appropriate plug-in components forms an application for certain neuropsychological rehabilitation tasks. If developers document the programming interfaces of their plug-in components as carefully as the NeurOSS platform will be documented, this will create interesting possibilities for component reuse in various other applications.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 740–743, 2006.

2 Open Source Software, Community Aspects and the 3rd World

The Open Source Initiative (OSI) [2] defines open source this way: “The basic idea behind open source is very simple: When programmers can read, redistribute, and modify the source code for a piece of software, the software evolves. People improve it, people adapt it, people fix bugs. And this can happen at a speed that, if one is used to the slow pace of conventional software development, seems astonishing.” It is up to the community how the NeurOSS project will succeed. If a fair amount of the projects creating software for neuropsychological tasks share their components for common use, the platform will reach the state of real importance to neuropsychological rehabilitation.

The idea of Open Source Software is especially tempting when one considers the health care sector where a large amount of research results are available and waiting to be implemented for the real-world usage. Most of the research is universal, which means that the results are equally valid everywhere. In developing countries [1] health care system has very limited resources for purchasing software products abroad, for example. However, many of those countries have succeeded in creating a working education system that trains highly skilled professionals. They could easily localize and adapt free open source products to local needs — if only there were such software products available.

3 Technical Overview

The NeurOSS could be described technically as a plug-in platform offering well-designed core services for neuropsychological rehabilitation software development, which enables easy integration of new innovative modules to the system. The NeurOSS system uses controlled blackboard approach [3] for interprocess communication, which allows plug-in programmers to easily pass data to other components of the system. However, to improve the efficiency of the system, the direct parameter passing from a component to another is possible. Direct signaling can be used as well. The first version of the NeurOSS platform has been written using Java programming language [4,5]. It allows also plug-ins, written by other programming languages, by using wrappers [5]. Java programming language was selected because it is highly platform independent and widely known by developer community. Additionally, high quality development tools are available for free. NeurOSS servers will offer limited possibility for distributed processing which means that NeurOSS servers can access each other´s blackboard and servers can send signaling messages with each other.

4 Plug-In Architechture

It is not a new idea to approach challenges of expandability by using plug-in architecture. The key idea of such architectures is to provide an organized

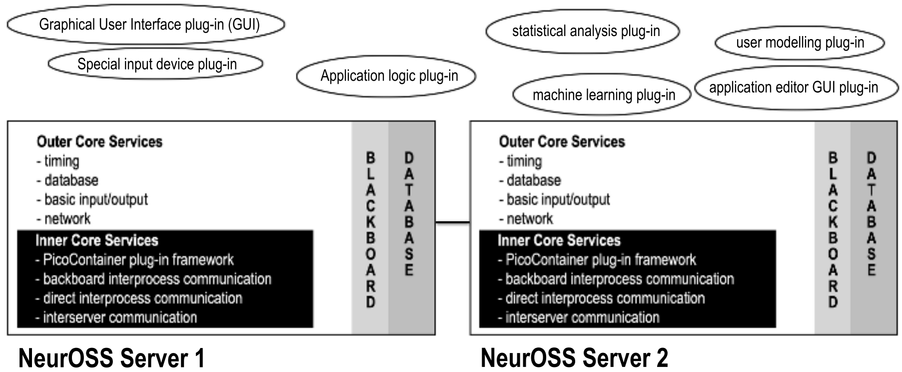


Fig. 1. Interconnected NeurOSS plug-in servers and some example plug-ins

way for developers to add new functionality to platform without knowing too much about the internals of the platform. There are quite a few existing plug-in frameworks available, e.g. Apache containers [6] and Eclipse [7] development environment. In the NeurOSS project we have selected to use widely known plug-in framework PicoContainer [8], because the framework makes it easy to modularize how dependencies between parts of an application are laced up by dependency injection [9]. It does not require a developer to implement any special APIs but just ordinary Plain Old Java Objects [10].

5 Platform Services

The inner core of the NeurOSS platform contains only lightweight memory management routines offering methods related to blackboard [3] which is used for interprocess communication, PicoContainer [8] framework handling plug-ins and their dependencies, and routines for communicating with other NeurOSS servers.

The outer core of the platform offers some content related services, so essential for neuropsychological rehabilitation applications that it has, for efficiency reasons, been reasonable to implement the functionality as core services. This level of services contains e.g. routines handling external database access, timing, basic mass-storage operations, basic input/output routines and simple network related routines.

All other functionality will be implemented as plug-in components if there is no clear reason to do otherwise. Everything related to application login, adaptivity, statistics, data analysis, and input and output routines for special devices, for example, will be implemented as plug-ins.

If there is a need for real-time user interfaces requiring very precise control of time, such components should be created as separate applications using only services offered by the NeurOSS platform. This is because real-time features of Java programming language are not included to the Java Standard Edition, which can be downloaded free of charge.

6 Discussion

As careful reader might notice, the architecture of the NeurOSS platform is equally well suitable for other kinds of application areas, not only for neuropsychological rehabilitation. The reason for narrowing the application area to neuropsychology relates mainly to the question of critical mass. As already mentioned, the main key for success is community. If the developer community can create enough open source components for different tasks, it will be possible to reach the critical amount of usefulness, which also draws other projects to use the NeurOSS platform. Project´s homepage at http://www.neuross.net/ will be opened by May 2006. At the first phase the homepage will contain latest information about the development of the platform and a discussion forum.

References

1. Rajani, N., Rekola, J., Mielonen, T.: FLOSS Report: Free as in Education: Significance of the Free/Libre and Open Source Software for Developing Countries (2003). http://www.maailma.kaapeli.fi/FLOSSReport1.0.htm
2. The Open Source Initiative (OSI): http://www.opensource.org/
3. Corkill, D.: Blackboard Systems. AI Expert 6(9):40-47 (1991). http://www.bbtech.org/papers/ai-expert.pdf
4. Java Programming Language: http://www.java.com/
5. Arnold, K., Gosling, J., Holmes, D.: Java Programming Language. 3rd edition. Addison Wesley Professional, (2000).
6. Apache Excalibur Project: http://excalibur.apache.org/
7. Eclipse Platform: http://www.eclipse.org/platform/
8. PicoContainer: http://www.picocontainer.org/
9. Fowler, M.: Dependency Injection (2004): http://martinfowler.com/articles/injection.html
10. Plain Old Java Object: http://en.wikipedia.org/wiki/Plain_Old_Java_Object

DALMA – Location Aware Alarm System for People with Disabilities

Roberto Casas, Álvaro Marco, Jorge L. Falcó,
Héctor Gracia, and José I. Artigas

TecnoDiscap Group. University of Zaragoza,
50018 Zaragoza, Spain
{rcasas, amarco, jfalco, hgracia, jiartigas}@unizar.es

Abstract. This paper presents a location aware alarm system developed to give each person greater and safer mobility by automatically detecting risk situations. Location awareness has been possible thanks to an indoors positioning system (IPS) based on Bluetooth and ultrasounds, developed by the authors. The IPS is able to give accuracies of several centimetres, using a reduced infrastructure (6 fixed beacons every 100 m^2). The alarm system has been implemented in several stances of a special education school. Preliminary evaluation results show its reliability and usability.

1 Introduction

Life expectancy is continuously growing, the same as the probability of survival after accidents or diseases. In both cases, people usually decrease their capacities becoming more dependent. If we add the disintegration of the traditional family structure in many countries, this leads to more people with disabilities living alone, in residences, or simply less attended. Adapting the environment to capacities of the individuals is being demanded from the society. Most people want to keep their independence and stay at home the longer time possible instead of living in residences [1]. Another important point is the thrift that this care inside the community allows to both, the official institutions and the person itself and his/her family [2].

Alarm systems have become key pieces inside assistive technology, because they enable people to extend the time living alone at home. Alarms increase the personal security of the individuals warning about several situations. Thus, people's quality of life is improved because their self-confidence increases and the anguish to be unattended diminishes; people know that in case of a crisis situation, they can contact immediately and be attended by someone able to manage the emergency situation [3].

Awareness of their possible disadvantages is also important. Isolation from the society of the individuals is the biggest risk, because of being more independent, people are less visited and attended. In many cases, a big part of the social contact of many elderly and people with disabilities is reduced to the caretaker

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 744–751, 2006.

who periodically visits them; here, the reduction of social contact can make alarm systems unadvisable.

This paper is organized as follows. First, the basics of alarm systems and how they can be enhanced through location is discussed. Then DALMA, a location aware alarm system, is presented detailing its insides, how some privacy issues have been implemented and its evaluation with real users. Finally, some conclusions are exposed.

2 Alarm Systems

The main objective of alarm systems is to inform the adequate entity that something is going wrong. This involves detecting the situation and giving a warning to get the attention of the responsible person or system. With regards to the detection, it can be passive —activated by the user— or active. First ones are widely used today; they are usually enabled by pressing a button in a pendant. Lamentably, there are many situations where these alarms are unusable: when the person does not have the capacity to recognize the emergency condition (e.g. mental disorientation access), when is not able to press the button (e.g. fall or unconsciousness) or just when the person never wants to disturb the caretaker [2].

Active alarms rely on technology to detect the emergency situations overcoming human factors. The habitual fire or gas leakage detectors would be in this category. More sophisticated sensors able to monitor bio-signals (pulse, ECG, skin impedance, etc.) are very useful to detect faints, infarcts and other health problems. Many authors also propose different devices to detect falls, overriding one of the biggest fears of the elderly [4,5]. Ambient intelligence extends the perceiving capabilities of the systems one step further than just using sensors. Cuddihy proposes merging information from motion, door and window sensors to detect activity patterns and alert the family when the home is unusually quiet [6]. In the same line, BT Exact is developing sensors to monitor activities such as preparing food, sleeping, playing, etc. [7].

Once the warning is generated, the system can head it to the user, caretaker, relatives or itself. In any case, its purpose is to reach an entity who can execute the actions needed to correct the alarm situation. The warning is sent to the user when he/she is presumably unaware of a detected risk, for example a gas leakage. Obviously this would only be applicable when the user is able to understand the information and act consequently: decide what to do or call another person. This is the most desirable situation because the user decides on his/her privacy every moment. Warning the caretaker or a relative is more common, it happens due to the user's incapacity to decide, urgency of the necessity, or because a chained call takes place (system warns the user and the user calls the caretaker). Warning entities external to the system as user or caretaker may reduce effectiveness, needed in emergencies. In some circumstances the system acts first (cuts the gas supply) and then warns the user or the caretaker if the user does not respond. Totally delegate in a machine, actions that affect a person, is not recommended. Nevertheless, machines are good executing support actions in the resolution of the warning and performing tasks of feedback to the user; for example informing a person that has fallen that the alarm situation has been detected and the caretaker has received the warning, i.e. somebody is coming.

Location awareness inside an alarm system increases the sort of situations detected. Detection of people lost or in dangerous places is the most evident application; there are commercial devices that integrate GPS in a mobile phone in order to warn the family when a person gets lost [8]. When moving into a private dwelling, location with door sensors, movement detectors or pressure mats is used to detect wanderings at night and stances in places where the person may be at risk [2]. All these applications, do not need accurate location nor identification of the person tracked. With an affordable system able to give this information indoors, the possibilities dramatically increase. Analyzing a person's movements can be useful for the identification of precursory patterns of an alarm situation. For example, if we determine that a person starts to describe circular trajectories before having an anxiety crisis, we could act on the crisis before it starts. Of course, merging this information with other obtained for example from bio-sensors, greatly empowers the detection capacities. If the positioning system is also able to locate multiple users, it can be used in scenarios such as residences where each person has different needs. Not only it enables detecting situations such as escapism or night wandering (big concerns in that places), it also informs about the situation of the person in case of another alarm occurs. For example, instead of solely detecting that a person has fallen down, it is possible to determine his/her position in order to provide faster assistance. All this results in an enlargement of the user's personal independence; people can move freely in a greater area being sure that they will be quickly found when an alarm occurs. One example of this would be a confused person walking in the gardens without companion. In principle, risk of disorientation, escapism or falls exists, but thanks to monitoring, these situations are detected and solved.

3 DALMA System

DALMA is a location aware alarm system; it senses several parameters such as temperature or falls and combines them with indoor positioning of people to detect various risk situations. DALMA shows its usefulness in big environments with several users having different disabilities, i.e. residences, hospitals and public or private centres. The positioning system we have developed is able to precisely locate and identify a person. As every user has different circumstances, the system personalizes the data analysis being able to detect the following location alarms:

– Risk due to stance in a room that supposes risk for that person without his/her caretaker (kitchen, healing room, etc.).
– Stance or wandering to be corrected (nocturnal stance in the dining room).
– Stance duration greater than the specified one (staying in the bath longer than half an hour can indicate that he/she has undergone some type of accident or that simply is disoriented and does not know how to leave the room).
– Absence or contact loss (the user is leaving or has left the building).

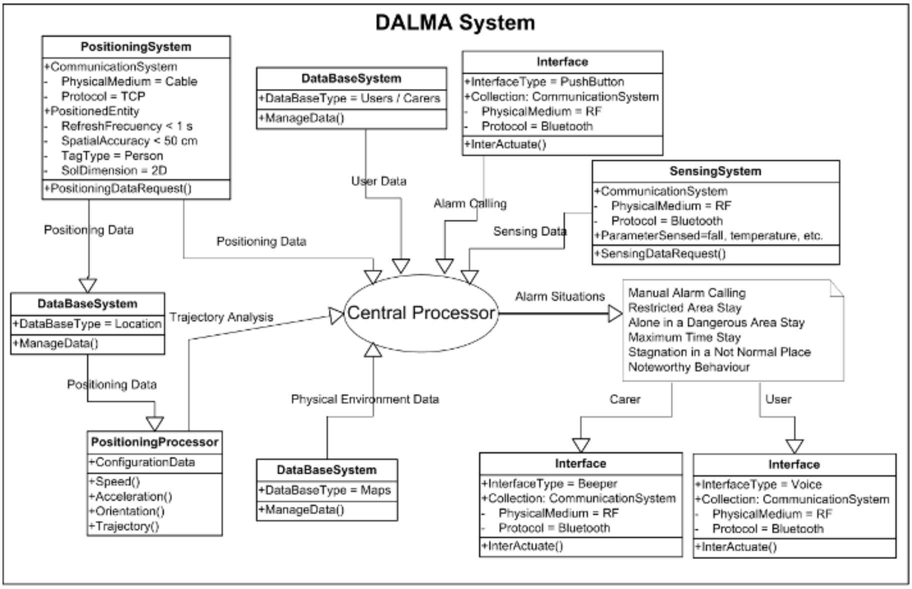


Fig. 1. UML model of DALMA system

– Conducts that denote anxiety, escapism or other pathologies (small repetitive itineraries).
– No movement for too long in a nonhabitual place (no movement for five minutes in the middle of the corridor).

Upon the detection of any kind of alarm —location, fall, button pulsation, etc.—, as the system knows where it has happened, it will trigger a warning to the most appropriate person able to solve it. The notified person will depend on the sort of situation detected and on the person itself. For example, if fast assistance is required, i.e. because of a fall or an escape attempt, the system will warn the nearest caretaker (or caretakers). It can go to the person itself that can overcome the situation because he/she just did not realize about it. In other cases, the person may have a favourite carer or, because of his/her characteristics, need the assistance of various at the same time; in all cases, the system will analyse the situation and personal preferences to solve the situation the best possible way.

3.1 DALMA Insides

DALMA's core is a central processor that performs an elaborated analysis of data coming from several sensing systems and databases to detect alarm situations which are solved warning the appropriate people. Fig. 1 represents a UML model of DALMA showing all the subsystems inside.

We have developed an indoor positioning system that works with Bluetooth and ultrasounds: BLUPS [9,10]. It consists of several beacons fixed in the ceiling and mobile tags to be located. The beacons successively broadcast ultrasonic pulses that are received by the tags, which measure the ultrasound time of flight. Every tag sends its identity and the time of flight data to a central control (personal computer) via Bluetooth, to compute the tag position. BLUPS is able to locate up to 100 tags, with an accuracy of 5 cm, and a maximum temporal resolution of one new location every second. The beacons have been implemented with the USI UB1-1112 Bluetooth module and a PIC 16F73 microcontroller that generates the 40 kHz signals for the ultrasound driver and manages the Bluetooth module using a simplified Host Controller Interface.

All this information is used in two ways: it is stored in the location database and sent to the central processor via TCP. The database is used by the positioning processor that computes the trajectories of people and, when matching a predefined one, informs the central processor. It can also compute speed, acceleration or orientation if they would be needed.

Other inputs to the system are those coming from several sensors. Our implementation can measure the ambient temperature (to detect extreme temperatures), sense falls with an accelerometer and capture keystrokes to ask for assistance. All this information, when demanded, is sent via Bluetooth to the central processor. Fig. 2 shows a photograph of the implemented tag for location, sensing and Bluetooth communication.

The central processor has also access to databases of all the users to check the situations to be detected and the preferences in case of alarm. It also uses a map database containing the information about the covered area, needed because the location system gives the position in absolute coordinates and the alarms are specified in terms of rooms.

When any kind of alarm is detected, the central processor determines who should be notified. This notification is made via an interface; in the caretakers' case, is something similar to a beeper with an LCD display where a description of the alarm appears. In case the destination is the person itself (or another companion that could help), we are working in a voice interface that allows to establish a bidirectional communication.

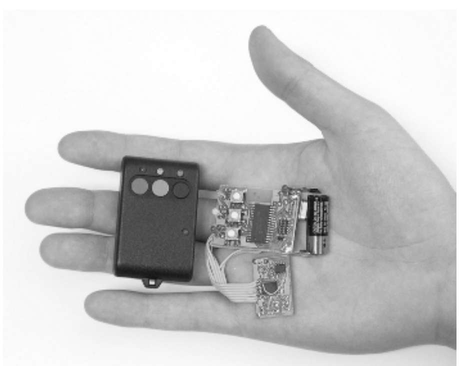


Fig. 2. Sensing, location and communications tag

3.2 Privacy in DALMA

Location of people not only has a big impact on the individual's privacy rights, it can also have very important psychological effects. If the location conditions are not clearly stated, it is easy to think we are being spied at any moment, with the stress this can suppose.

The use of a positioning system entails an obvious utility for them. Nevertheless, due to their special characteristics, a reflection becomes necessary on the ethical aspects related to the invasion of the private life and the restriction of freedom that it could suppose. Considering the abilities of the person, we can distinguish two different assumptions. If the person is able to understand the situation and show her/his consent, he/she accepts or rejects the service [11]. In this case the ethical requirements refer to the treatment given to the data issued by the system and the validity of the way of informing the user to ask his/her consent. It would be necessary to establish a protocol detailing the information to be gathered, the use planned for it, how and how long it is going to be stored or destroyed, who will have access to it, etc. This protocol would have to be explained to the user clearly in order to have his/her informed consent.

When the person is not able to declare his/her consent, a conflict between two rights is established: in the one hand the right to the intimacy and by another one the right to the health and even to the personal well-being [12]. This could be solved, like it happens in the medical ambit, prioritizing the protection of the person over the privacy right. Although we are always looking to protect the person, it would be necessary that the one holding the legal guardianship on the person with disability gives his/her consent. In the OMA guidelines an analogue problem is stated: the tracking of the under aged children by their parents. They indicate that only the informed consent of legal tutor is necessary, but recommend at least the notification to the minor [11].

Applying the above privacy concepts to DALMA, we distinguish different user profiles from the stakeholders' point of view. The caretakers, as employees, could be located provided a signed explicit consent [13]. However, in order to reduce the stress caused by the "big brother" effect, they are only located when an alarm occurs. People having sporadic crisis that can be detected by the system or people with accident risks (elderly) are only located when the system detects an alarm situation (fall, disorientation...) or when they request help voluntarily. In this case the positioning information is not stored; it is used solely to deal with the situations effectively. Finally we have the people who are periodically positioned in order to detect risk situations (repetitive itineraries, stances in dangerous places...). Usually this people will not be able to give an informed consent, so their legal tutor should sign it, but always paying attention to protect them. In this case, the location information could be stored the time necessary to be analyzed and then it should be destroyed.

The technology provides adequate security in data management and communications. Concretely, as the location system is centralized, we use database encryption, security access controlled by passwords and pseudonyms [14] to help on keeping the anonymity of the people.

3.3 System Evaluation

DALMA has been installed in a special education school in Spain, Los Pueyos. It covered some classrooms, corridors, the office and the gymnasium. Ten students carried the tags while developing their normal activities and two caretakers had a beeper. The system perfectly tracked all the students, warning when entering the office without any caretaker and staying in the corridor for more than three minutes (the programmed warning situations). It also detected falls (they were simulated in the gymnasium) and keystrokes. Depending on the user, the warnings went to his/her favourite caretaker except when fallings. The interface was considered good enough by the caretakers, but the messages and information appearing in the displays should be enhanced. The analysis of trajectories was not enough to correctly identify them, now we are developing an improved analyzer using neuronal networks.

The global evaluation of the system is good, but it is needed to test it in a bigger scenario. Thus, we are going to install it in a residence for people with disabilities that is currently under construction.

4 Conclusion

This paper has presented DALMA, a location aware alarm system specially developed for elderly or disabled people. Its main aim is giving each person greater and safer mobility by automatically detecting risk situations. Thanks to it, the personal environments can be enlarged; it is possible to diminish coercive physical barriers (locked doors, closed dependencies, etc.) while allowing the person to make activities that may have a risk which can be detected through monitoring. The system extends the individual's capacity and independence and reduces the time of suffering due to long delays waiting for assistance in case of accident. In many cases the system can also suppose a psychological reinforcement for the person, who feels safer to undertake actions alone. From the caretakers' point of view, this system alleviates the stress they usually feel when caring several people at the same time.

Acknowledgments

This work was supported in part by the Spanish MCYT under Project FIT-150200-2000-215 and by DGA under project T44.

References

1. Dewsbury, G., Edge, M.: Designing the Home to Meet the Needs of Tomorrow... Today: Smart Technology, Health and Well-being. Open House International, Vol. 26. (2001) 33–42.
2. Miskelly, F.G.: Assistive technology in elderly care. Age and Ageing, Vol. 3. Oxford Journals (2001) 455–458.

3. Calling for Help Group: Community Alarms Services: A National Survey. London Research Institute for Consumer Affairs (1994).
4. Sixsmith, A., Johnson, N.: A Smart Sensor to Detect the Falls of the Elderly. IEEE Pervasive Computing, Vol. 3 (2004) 42–47.
5. Riisgaard, T., Hansen, J., Eklund, M., Sprinkle, J., Bajcsy, R., Sastry, S.: Using smart sensors and a camera phone to detect and verify the fall of elderly persons. European Medical and Biological Engineering Conference (2005).
6. Cuddihy, P.: Home Assurance System, IEEE Pervasive Computing, Vol. 3 (2004) 48.
7. Brown, S.J.: Next-generation Telecare System. IEEE Pervasive Computing, Vol. 3 (2004) 50.
8. Benefon Company. http://www.benefon.com/catalogue/gsm_gps/.
9. Casas, R., Marco, A., Guerrero, J.J., Falcó, J.: Robust Estimator for Non-Line-of-Sight Error Mitigation in Indoor Localization, EURASIP Journal on Applied Signal Processing, (2006).
10. Sevillano, J.L., Falcó, J., Abascal, J., Civit-Balcells, A., Jiménez, G., Vicente, S., Casas, R.: On the Design of Ambient Intelligent Systems in the Context of Assistive Technologies, Lecture Notes in Computer Science, vol. 3118 (2004) 914–921.
11. Location working group of Open Mobile Alliance (OMA): Privacy Guidelines. (2002).
12. Pompano, J.R.: Privacy Issues and Disabled Persons. Yale-New Haven Teachers Institute, (2000).
13. Langheinrich, M.: Privacy by Design – Principles of Privacy Aware Ubiquitous Systems. International conference on Ubiquitous Computing (2001) 273–291.
14. Beresford, A., Stajano, F.: Location Privacy in Pervasive Computing, IEEE Pervasive Computing, vol. 2, n° 1 (2003) 46–55.

Usage of IT and Electronic Devices, and Its Structure, for Community-Dwelling Elderly

Madoka Ogawa[1], Hiroki Inagaki[2], and Yasuyuki Gondo[2]

[1] Obirin University Graduate School,
2-9-2 Yoyogi, Shibuya-ku, Tokyo, 151-0053 Japan
s0442603@obirin.ac.jp
[2] Tokyo Metropolitan Institute of Gerontology,
35-2 Sakae-cho, Itabashi-ku, Tokyo, 173-0015 Japan
{inagaki, ygondo}@tmig.or.jp

Abstract. Electrical household appliances and IT (information technology) are believed to increase the QOL and well-being of the people who use them. The benefits of electronic devices for elderly people would be more evident than for younger people because it is assumed that such equipment would compensate for the decline of functional ability in the elderly. However, there has been only very limited research on the actual usage and influence of such devices in relation to generation and age. The purposes of the present study were to clarify the actual situation with regard to the use of IT and electronic devices by community-dwelling elderly, and to characterize individuals according to their familiarity with such devices.

1 Introduction

The use of IT and electronic devices in daily life is now commonplace and has brought various benefits [1,2]. For example, household appliances reduce the time spent on housework, and create time for leisure activity. In particular, popularization of the Internet has revolutionized accessibility to various sources of information and offers a wide range of choices from multiple options [1,3]. The popularization of these devices has undoubtedly been due to the resulting contribution they make to the quality of life (QOL) and well-being of the users. It is expected that the remarkable progress currently being made in IT and electronic devices will bring further benefits to society [1].

Recently, however, the issue of the “digital divide” between members of society who have benefited from advances in IT and those who have not, has begun to receive attention. Elderly people are considered to represent a section of society who have been bypassed by the IT revolution, most likely because of their reduced peripheral sensory and cognitive function, and lack of mobility or sufficient literacy or familiarity with IT and electronic devices [4,5]. Developed countries are now heading towards an aged society in which people of very advanced age will not be exceptional. Information technology changes at a rapid pace, but the digital divide still exists [3]. Therefore, it is expected that the problem of the "digital divide" in the elderly population will become more acute.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 752 – 758, 2006.

Our research has been aimed at developing a program for promoting the use of new devices by the elderly. Two issues are intrinsic to this aim. The first is to improve devices such as interfaces, and the second is to remove environmental and psychological barriers. This report addresses the latter issue. Here we describe, firstly, the actual situation with regard to the use of IT and electronic devices by community-dwelling elderly, and secondly, we characterize individuals in terms of their use of IT and electronic devices. We believe this work will help to clarify the factors that promote or inhibit the use of IT and electronic devices by the elderly.

2 Method

2.1 Participants

This study was carried out as part of a longitudinal follow-up project to investigate the impact of life events on the psychological well-being of elderly individuals. In the first year (1991), 3,097 community-dwelling middle-aged to elderly people (1,362 males and 1,735 females, aged 50-74 yr) participated. A follow-up survey was carried out every year until 2000 [6]. At the last examination in 2005, 1,169 (456 males and 713 females, aged 65-89 yr) individuals participated. We used the data for 1,121 people (431 males and 690 females) with a mean age of 73.7±6.2 yr (male: 73.5±6.1 yr, female: 73.7±6.3 yr). Table 1 shows the background characteristics of the participants.

2.2 Procedure

We used two discrete research methods – a visit survey and an invitation survey – according to the participants' age and preference. Basically, we interviewed the participants themselves. Each individual was interviewed by a trained university student who

Table 1. Background characteristics of participants

	N	%	%men
Age group			
65-69	330	29.4	37.6
70-74	314	28.0	39.5
75-79	266	23.7	41.4
80-84	152	13.6	31.6
85-89	59	5.3	42.4
Education			
no education	46	4.1	34.8
elementary education	295	26.3	34.6
secondary education	485	43.3	30.5
higher education	270	24.1	57.4
unknown	25	2.2	40.0

was studying psychology or gerontology and trained adult interviewer. The interview took about one and a half hours. The study was approved by the ethics committee of Tokyo Metropolitan Institute of Gerontology and each participant gave written informed consent.

2.3 Measurement

In addition to the use of a questionnaire on physical function, psychological function and cognitive function, we asked the participants about the frequency of use and degree to which they had mastered the operation procedure for 21 items of IT and electronic devices. The use frequency and degree were expressed as one of four choices: 1. "I think I use it well and have mastered it", 2. "I use it, but only its fundamental function", 3. "I've used it before, but have stopped using it now", and 4. "I've never used it before". Table 2 lists these 21 devices and the percentages of individuals who use them (who answered 1 or 2 to the above question). Furthermore, we asked about the reasons for use or non-use of individual devices.

3 Results

Table 2 shows the average frequency of use of each device. As can be seen, "analog TV" (98.8%) and "air-conditioning equipment" (91.0%) had a high frequency of use, and "L-mode" (2.1%) and "Internet bank" (1.8%) had a low frequency of use. Because of their low frequency, we excluded the latter two items from further analysis.

For the first analysis, to classify IT and electronic devices, we carried out a hierarchical cluster analysis of the 19 items. As a result, three categories emerged: "IT and digital devices" (digital TV, DVD player, video camera, digital camera, personal computer, word processor, Internet), "Classical household appliances" (microwave oven, automatic washing machine, rice cooker, air-conditioning equipment, analog TV), "Quasi-household appliances+ATM" (radio, component stereo set, video tape recorder/HDD recorder/DVD recorder, answering machine, fax, cellular phone, ATM)(Table 2).

We used two-way ANOVA to compare the frequency of use in all items and each of the three categories for sex and age group (65-69 yr, 70-74 yr, 75-79 yr, 80-84 yr and 85-89 yr) as independent variables. As a result, there was a significant main effect of sex and age group in terms of all items and the three categories. Further multiple comparisons indicated that the critical age at which a decrease in use became evident differed among the categories and sexes. As a whole, older age groups and women used devices significantly less frequently than younger age groups and men, and the trend of the decrease was linear. For "IT and digital devices", a linear decrease according to age was seen for men, while a large difference was observed between the ages of 60 years and 70 years for women. For "Classical household appliances" the decreasing trend was the same in total for both men and women. For "Quasi-household appliances+ATM", a large difference was observed between individuals aged less than 75 years and those aged 75 years and older, while a linear decrease with age was seen for women. Figure 1 shows the percentage of device use by sex and age group.

Table 2. Three clusters of devices and their total percentage of use

items	total % of use
IT and digital devices	
digital camera	18.7
digital TV	17.4
DVD player	17.2
personal computer	16.8
word processor	16.6
internet	11.8
video camera	10.1
Classical household appliances	
analog TV	98.8
air-conditioning equipment	91.0
microwave oven	89.3
rice cooker	83.2
automatic washing machine	70.3
Quasi-household appliances+ATM	
answering machine	66.4
ATM	65.3
radio	61.1
component stereo set	45.8
video tape recorder/HDD recorder/DVD recorder	40.9
fax	39.4
cellular phone	37.4
Other items	
L-mode (e-mail and information service by analog telephone)	2.1
Internet bank	1.8

For the second analysis, we classified the participants by hierarchical cluster analysis according to use of the 19 devices. As a result, five groups emerged. Table 3 shows the mean age, sex and average number of devices used in each group. To describe the characteristics of those five groups, we compared the use of the devices among the groups. One-way ANOVA revealed the main effect in each group category, and further multiple comparisons in each device category showed that "All items" and "IT and digital devices" were used markedly less frequently by groups 4 and 5 than by the other three groups. "Classical household appliances" were used markedly less frequently by group 5 than by the other four groups. "Quasi-household appliances+ATM" were used markedly less frequently by groups 3, 4, and 5 than by the other two groups.

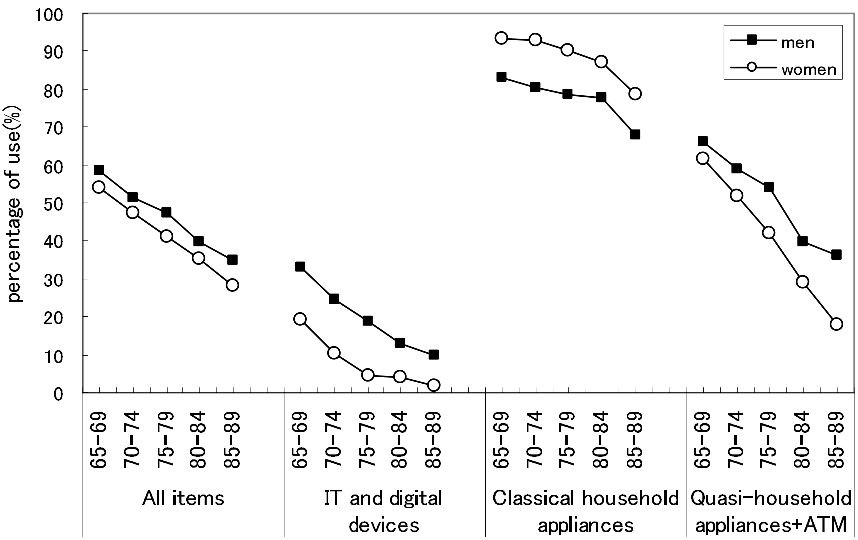


Fig. 1. Percentage of device use by sex and age group

We named these five groups "advanced", "developing", "stagnation", "conservative" and "non-use", respectively, based on the characteristics of their device use. The characteristics of these group are as follows, the first group ("advanced") comprised individuals who used all three categories, had the lowest average age, and had a high male ratio. The second group ("developing") comprised individuals who used "Classical household appliances" and "Quasi- household appliances+ATM", had a comparatively low average age, and a roughly equal male:female ratio. The third group ("stagnation") comprised individuals who used mainly "Classical household appliances" but used "Quasi-household appliances+ATM" to a low degree, had a comparatively low average age, and were predominantly female. The fourth group ("conservative") comprised individuals who used only "Classical household appliances", had a comparatively high average age, and were predominantly female. The fifth group ("non-use") comprised individuals who did not use any IT and electronic devices, had the highest average age, and showed a roughly equal male:female ratio.

4 Discussion

The purpose of this study was to investigate elderly people in terms of their use of IT and electronic devices and to clarify their characteristics in order to provide basic data for developing a program to promote the use of such devices. We interviewed 1,121 community-dwelling elderly people and succeeded in classifying both the devices and the individuals who used them into distinct groups.

Table 3. Background and means of number of use, mean age and sex ratio as a factor of categorized use grouping

	1 advanced	2 developing	3 stagnation	4 conservative	5 non-use
Mean age	69.7	71.6	73.1	76.0	76.5
Sex (%men)	68.3	41.9	34.0	20.7	50.5
Average number of use					
All items	14.9	12.3	9.0	6.3	5.2
IT and digital devices	4.9	1.7	0.5	0.1	0.3
Classical Household appliances	4.3	4.7	4.6	4.6	3.0
Quasi-household appliances+ATM	5.7	5.8	3.9	1.6	2.0

First, the IT and electronic devices were classified into three categories: “IT and digital devices”, “Classical household appliances” and “Quasi- household appliances+ATM”. “IT and digital devices” included digital appliances, such as the PC and Internet, which have appeared on the market in recent years. “Classical household appliances” included household electrical appliances with high diffusion rates that have been commonly use for many years, such as the microwave oven and analog TV. “Quasi-household appliances+ATM” included home electronic appliances, such as the video and answering machine, which have already been in use for a relatively long period. We speculated that the combination of subject age and the diffusion rate of a given device would have a strong influence on whether the device was used or not used. For example the diffusion rate of the TV, which is a representative “Classical household appliance”, surpassed 50% as long ago as 1972. On the other hand, the diffusion rate of the VTR, which is a representative of “Quasi-household appliances+ATM”, did not surpass 50% until 1988. The individuals in the oldest age group (85+) would have been in their late 50s during the early 1970s and they were almost 70s in the late 1980s. This age difference would have influenced both their motivation as well as their competency to use the device. Detailed analysis of the relationship between the diffusion rate for each item and the age of individuals who use it will be required.

Secondly, we succeeded in classifying the participants into five groups: “advanced”, “developing”, “stagnation”, “conservative” and “non-use”. A simple analysis indicated that age and sex were strongly associated with each group category. The present study did not include analysis focusing on functional status such as vision, hearing and mobility, or cognitive status. In addition, we hypothesized that other factors such as educational attainment, job history, living situation and personality would also influence the use or non-use of IT and electronics devices. To clarify the detailed pathway determining use or non-use, we are planning to perform joint analysis incorporating previously corrected longitudinal datasets.

References

1. Wahl, H.W. and Mollenkopf, H.: Impact of everyday technology in the home environment on older adults' quality of life. In: Charness, N. and Schaie, K.W. (eds.): Impact of technology on successful aging. Springer, New York (2003) 215-241
2. Park, D.C. and Jones, T.R.: Medication adherence and aging. In: Fisk, A.D. and Rogers, W.A. (eds.): Handbook of human factors and the older adult. Academic Press, San Diego (1997) 257-288
3. Kiel, J.M.: The digital divide; Internet and e-mail use by the elderly. Med Inform Internet Med 30 (2005) 19-23
4. Howard, J. and Howard, D.: Learning and memory. In: Fisk, A.D. and Rogers, W.A. (eds.): Handbook of human factors and the older adult. Academic Press, San Diego (1997) 7-26
5. Fisk, A.D. and Rogers, W.A.(eds.): Handbook of human factors and the older adult. Academic Press, San Diego (1997)
6. Shibata, H., Suzuki, T. and Shimonaka, Y.: Overview of a new longitudinal interdisciplinary study on aging (TMIG-LISA,1991-2001). Longitudinal Interdisciplinary Study on Aging 2nd edition. Serdi Publisher, France (1997) 7-20

A Randomized, Placebo-Controlled Study of the Efficacy of Cognitive Intervention on Elderly People and on Patient's with Alzheimer's Disease

J.J. Yanguas, C. Buiza, I. Etxeberria, N. Galdona,
M.F. González, and E. Urdaneta

San Sebastian, Spain
jyanguas@fmatia.net

Abstract. Research on non-pharmacological therapies (cognitive rehabilitation) in old age has been very limited, and most has not considered the effect of intervention of this type over extended periods of time. The aim of Donostia Longitudinal Study was to investigate a new cognitive therapy in a randomized, placebo-controlled group of elderly people over 65 years of age without cognitive deterioration or expressed AAMI (Age Associated Memory Impairment) and patients with a Alzheimer's Disease (AD). The efficacy of this therapy was evaluated by means of post-hoc analysis of 390 people using biomedical, neuropsy-chological, affective, and personality assessments. In the GDS 1-2 grouping, scores for learning potential and different types of memory for the treatment group improved significantly relative to the untreated controls. While, subjects with a GDS 3-4 showed significantly better performance on Neuropsychiatric inventory (NPI) scores in all domains (anxiety, depression, apathy, sleep disturbances). Finally subjects with a GDS 5-6, showed a maintenance of cognitive capacities. In the GDS 1-2 grouping, the most significant result found is that learning potential of trained people enhances within two years of intervention, this involves a successful ageing sign and plays a preventive role in dementia development. On the other hand, in the GDS 3-4 and GDS 5-6 grouping, the behavioral disturbances diminished within this intervention, so this type of training program could be beneficial on them.

1 Introduction

The rising population of older people is a fact of twenty-first century life. In Europe there were 74 million people over the age of 65 in 2003, compared with just 38 million in 1960. Today, senior citizens make up 16% of the total population, and this percentage is expected to climb to 27% in 2010 [1]. Although cognitive decline in advanced age has been recognized throughout history, the understanding that it represents the result of specific disease states is more recent. Longer life spans and increasing knowledge of the causes of cognitive decline, particularly Alzheimer's disease, has led to the prediction of dementia as an epidemic extending into the 21st century.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 759 – 765, 2006.

Prevalence rates of Alzheimer's Disease (AD) were calculated for the community population from the sample undergoing clinical evaluation. Of those over the age of 65 years, an estimated 10.3% (95% confidence limits, 8.1% and 12.5%) had probable Alzheimer's disease. This prevalence rate was strongly associated with age. Of those 65 to 74 years old, 3.0% (95% confidence limits, 0.8 and 5.2) had probable Alzheimer's disease, compared with 18.7% (95% confidence limits, 13.2 and 24.2) of those 75 to 84 years old and 47.2% (95% confidence limits, 37.0 and 63.2) of those over 85 years [2].

2 State of the Art

Numerous studies have demonstrated the important role played by cognitive training in averting age-associated memory alterations. Therefore, it is likely that brain plasticity is distinctly possible in the elderly [3]. The ACTIVE (Advanced Cognitive Training for Independent and Vital Elderly) group has published a longitudinal study on cognitive training and older people [4]. Their results showed that cognitive interventions helped older people to improve their cognitive yield (episodic memory, reasoning, and processing speed) over two years. Cappa et al. [5] studied the effectiveness of non-pharmacological treatments in improving older people's memory in the framework of the recommendations of the European Federation of Neurological Societies (EFNS), and they concluded that memory improved significantly following cognitive training.

The neuropsychological rehabilitation of patients with Alzheimer’s Disease (AD) have been focused in the development of neurostimulating programs that strengthen or maintain the nodeteriorated abilities. Davis, Massman and Doody [6] analyze the efficacy of a cognitive intervention consisting on training in face-name associations, spaced retrieval and cognitive stimulation. Thirty seven patients with AD received cognitive intervention for 5 weeks. Results suggested that face-name training, spaced retrieval and cognitive stimulation may produce small gains in learning personal information and on a measure of attention. Another study analyze the effect of cognitive stimulation training on older people with the diagnosis of dementia of Alzheimer’s type. The 78 community-dwelling patients were assessed on cognitive and behavioural functioning and randomly assigned to three conditions. The experimental group improved in cognitive and behavioral performance with treatment, but returned to former level functioning by the 9^{th} month.

Still, very few studies have maintained follow-up for more protracted periods and/or have employed sufficiently large samples to enable their findings to be viewed with confidence, and hence the results must be considered with caution [7].For all these reasons, the Donostia Longitudinal Study (DLS) was designed to determine the benefits of cognitive intervention in a large sample of subjects divided into three separate GDS (Global Deterioration Scale) groupings, namely, GDS 1-2 (normal aging), GDS 3-4 (mild-moderate AD), and GDS 5-6 (severe AD) [8].

The principal DLS objective was to study the effectiveness of non-pharmacological cognitive intervention in older people who either exhibited no cognitive deterioration or expressed AAMI (Age Associated Memory Impairment) and in patients with probable Alzheimer's disease. Effectiveness being measured as objective improvement in the scores for the cognitive functions tested, both after treatment and without treatment.

3 Methods

3.1 Design

The DLS employed a quasi-experimental, double-blind design that included different treatment conditions, namely, an experimental, a placebo, and a control group. In assigning subjects to the respective groups, gender, age, and cardiovascular risk were taken into account.

3.2 Subjects

Most of the 390 participants were recruited at various retirement homes in the Guipuzcoa region in Spain and delivered form the Neurology Service of the Basque Health Service. Written informed consent or witnessed oral consent to participate in the programme was obtained from each patient before the beginning of the study. The requirements for inclusion in the study were to be over 65 years of age without cognitive deterioration or expressed AAMI (Age Associated Memory impairment) and patients with a Alzheimer's Disease (AD). The main causes of exclusion from the study were: (a) any degenerative neurological disorder other than AD (Parkinson's disease, epilepsy, progressive supranuclear paralysis, subdural bruise or other injury, convulsive disorders, multiple sclerosis, etc.); (b) severe psychotic traits, depression, agitation or behavioural problems that might prevent successful completion of the programme; (c) a history of alcohol or substance abuse; (d) schizophrenia; (e) systemic disease estimated as being likely to yield a life expectancy of less than one year.

The total sample was then divided into three separate GDS (Global Deterioration Scale) groupings namely; GDS 1-2 (Normal aging. This group was divided in three treatment groupings: experimental, placebo, control), GDS 3-4 (Mild-moderate AD. This group was divided in three subgroups and GDS 5-6 (Severe AD. This group was divided in two treatment groups).

3.3 Programme

The Donostia Longitudinal Study (DLS) lasted two years, during which time each patient underwent six evaluations, an initial assessment followed by another every six months. This intervention grew out of a theoretical model developed by the authors based on Braak & Braak's model of Alzheimer's staging [9, 10, 11], which recognizes the existence of pathological anatomical correlates that affect different areas of the brain at the various stages of Alzheimer's disease (AD), with a specific disease course.

Training in the experimental and placebo groups took place in a group context. Session frequency was twice weekly and session length was one hour and a half in the GDS1-2 group. There were a total of 180 sessions during the two years of intervention some of them focused on cognitive training and others in health related issues. In GDS 3-4 group, the session's frequency was three days per week and lasted for two hours. In the experimental group a same function along sessions in a week were trained, with the objective to reinforce these functions. On the other hand, in a weekly structure, one of the three sessions was related with cognitive functioning, the second session was related to well-being or Activities of Daily Living and finally, the third session integrated the contents of the cognitive and functional capacities. Finally in GDS 5-6 group, the session frequency was four times per week and session duration was of thirty minutes. The session contents were focused on the cognitive stimulation on preserved cognitive functions and the maintaining of the Activities of Daily Living.

3.4 Measurement

Over the course of this study six assessments were carried out per subject. Of these, four followed an established protocol set out in Table 1 and 2, and two were control assessments using the ADAS-COG test [12]. Except in the GDS 5-6 (severe AD) group, where the control assessments were carried out using the SCIP test [13]. First, a multicomponent assessment was performed by an interdisciplinary team made up of a neurologist, four neuropsychologists, a gerontologist, and a social worker, all suitably experienced and specifically trained for this study. The assessment was approved by the Matia Foundation Ethics Committee. Only behavioral and cognitive evaluations are presented here.

Table 1. Behavioral assessment protocol

ASSESMENT	VARIABLE	TEST	REFERENCE
Behavioural Assessment	Behavioural disturbances	Neuropsychiatric inventory	[17]

3.5 Statistical Methods

Statistical analysis was carried out using the SPSS program, version 12.0 using a level of significance of less than 0.05. Values are the mean and the standard deviation. The Kruskal-Wallis test was used to establish differences within each group over time.

Table 2. Cognitive assessment protocol

Cognitive Assessment (GDS 1-2 and GDS 3-4)	Temporal, spatial and personal orientation	Information and orientation of WMS-R	[18]
	-Attention and working memory	Direct and inverse digits of WMS-R	[18]
	-Immediate execution memory -Logic memory	Logic memory of WMS-R	[18]
	-Recent word list memory -Short term memory -Learning potential	Auditory verbal learning test (AVLT)	[19]
	Designation language	Boston Vocabulary test	[20]
	-Repetition language -Audit Compression -Written compression -Written language -Reading language	-Boston Diagnostic Aphasia examination (BDAE)	[20]
	Visuconstructive ability	WaisIII(BLOCKS)	[21]
	Planning	Clock drawing (order and copy)	[22]
	-Bimanual coordination -Pre-motor function	Motor sequences of Luria	[23]
	-Visomanual coordination speed -Visomanual coordination execution	Trail Making Test, Part A	[24]
	Phonetic fluency Semantic fluency	FAS	[25]
	Abstraction	Proverbs	
	Categorization	Similarities of WAIS-III	[21]
	ICPR IMPR	Motor Sequences of Luria	[23]
Cognitive Assessment (GDS 5-6)	Cognitive functioning	SCIP-Severe cognitive impairment profile	[13]

4 Results

In the *GDS 1-2 grouping*, scores for learning potential and different types of memory (working memory, immediate memory and logic memory) for the treatment group improved significantly relative to the untreated controls. Subjects with a *GDS 3-4,* showed significantly (P<0.001) better performance on Neuropsychiatric inventory (NPI) scores in all domains (anxiety, depression, apathy, sleep disturbances) and premotor function. In addition, post hoc analyses demonstrated no changes of experimental over placebo group in cognitive functions. Finally *GDS 5-6* grouping, showed a maintenance of cognitive capacities of patients with severe cognitive impairment, the cognitive functions that obtained better results were attention, long term memory, language, visoconstructive capacity and categorization ability.

5 Conclusions

- In the GDS 1-2 grouping, the most significant result found is that learning potential of trained people enhances within two years of intervention, this involves a successful ageing sign and plays a preventive role in dementia development.
- Subjects with a GDS 3-4, the behavioural disturbances diminished within this intervention, so this type of training program could be beneficial on them.
- Finally, Subjects with GDS 5-6, the results show the effectiveness of this kind of interventions on patients with severe cognitive deterioration not only on cognitive functions also in functional capacities.

The information that we have get among the Donostia Longitudinal Study about the progressive loss of cognitive function caused by aging and Alzheimer´s disease and the effectiveness of non-pharmacological cognitive interventions in some of the cognitive functions, may be useful for the development of cognitive skills that will improve the well-being and independence of people suffering from Alzheimer´s disease In addition, this information will be useful for the development of specific tools such as interactive computer-based cognitive training, cognitive aids (reminders), smart phones offering services like shopping.

References

1. Eurostat yearbook 2004 The statistical guide to Europe - Data 1992-2002.
2. Evans, D.A., Funkestein, M.S., Albert, M.S., Scherr, P.A., Cook, N.R., Chown, M.J., Hebert, L.E., Hennekens, C.H., Taylor, J.O.. Prevalence of Alzheimer´s disease in a community population of older persons. Higher than previously reported. Jama 262 (I8) (1989)
3. Kopytova, F.V, Mednikova, Y.S, Popova, E.N. Age-related structural and functional characteristics of rabbit hippocampal neurons during the formation of temporal associations. Neurosci Behav Physiol 34 (2004), 889-896.
4. Ball, K., Berch, D., Helmers, K., Jobe, J., Leveck, M., Marsiske, M., Morris, J., Rebok, G., Smith, D., Tennstedt, S., Unverzagt, F., Willis, S. Effects of cognitive training interventions with older adults. *JAMA* 18 (2002), 2271-2281
5. Cappa, S.F., Benke, T., Clarke, S., Rossi, B., Stemmer, B., Heugten, C.M. Members of the Task Force on Cognitive Rehabilitation. EFNS guidelines on cognitive rehabilitation: report of an EFNS task force. Eur J Neurol 12 (2005), 665-680

6. Davis, R.N., Massman, P.J., Doody, R.S. Cognitive intervention in Alzheimer disease: a randomised placebo –controlled study. Alzheimer Dis. Assoc. Disord. 15(1) (2001), 1-9
7. Selwood, A., Thorgrimsen, L., Orrell, M. Quality of live in dementia. A one year follow-up study International Journal of Geriatric Psychiatry 20, 232-237 (2005)
8. Reisberg, B., Ferris, S.H., De Leon, M.J, Crook, T. The Global Deterioration Scale of Assessment of primary degenerative dementia. American Journal of Psychiatry 139 (1982), 1136-1139
9. Braak, H., Braak, E. Neuropathologycal staging of Alzheimer related changes. Acta neuropathologica 82 (1991), 239-259
10. Braak, H., Braak, E. Morphological criteria for the recognition of Alzheimer's disease and the distribution pattern of cortical changes related to this disorder. Neurobiol Aging 15 (1994), 355-356
11. Braak, H., Braak, E.. Staging of Alzheimer-related cortical destruction Int Psychogeriatr 9 (1997), 257-261
12. Rosen, W.G, Mohs, R.C, Davis, K.L. A new rating scale for Alzheimer´s disease. American Journal of Psychiatry 141**,** 1356-1364 (1984)
13. Peavy, G. M., Salmon, D. P., Rice, V. A., Butters, N. Assessment of severely demented elderly. Journal of the international Neuropsychological Society 1 (1995), 185
14. Hachinski, C.V, Lliff, L.D, Zilkha, E.. Cerebral blood flow of dementia. *Archives of Neurology* 32 (1975): 632-637
15. Mahoney, F.I., Barthel, D.. Functional evaluation: the Barthel index. Maryland State Medical Journal 14 (1965), 61-65
16. Lawton, M.P, Brody, E.M. Assesment of older people: Self-maintaining and instrumental activities of daily living. The Gerontologist 9 (1969), 179-186
17. Cummings, J. L. The Neuropsychiatric inventory: Comprehensive assessment of psycopathology in dementia. Neurology 44 (1994), 2308-2314
18. Wechsler, D. Wechsler memory scale revised manual. Psychological corporation: San Antonio (1987)
19. Rey, A. Léxamen clinique en psychologie. Presses universitaires de France: Paris (1964)
20. Goodglass, H., Kaplan, E. Test de vocabulario de Boston. Editorial médica panamericana: Madrid (1996)
21. Wechsler, D. Adaptación española de la Wechsler Adult Intelligence Scale-III. TEA Ediciones, SA: Madrid (1999)
22. Goodglass, H., Kaplan, E.. Evaluación de la afasia y de trastornos relacionados. Editorial Médica Panamericana: Madrid (1983)
23. Manga, D., Ramos, F. Secuencias motoras de Luria. Luria DNA. Diagnóstico neuro-psicológico de adultos. TEA ediciones: Madrid (2001)
24. Agree, E. & Freedman, V. Incorporating assistive devices into community-based long-term care: An analysis of the potential for substitution and supplementation. Journal of Aging and Health 12 (2000), 426-450
25. Benton, A.L., Hamsher, K. Controlled oral word association. Multilingual aphasia examination. AJA Associates, Iowa: USA (1989)
26. Taylor, D. & Hoenig, H. The effect of equipment usage and residual task difficulty on use of personal assistance, days in bed, and nursing home placement. Journal of the American Geriatrics Society 52 (2004), 72-79
27. Verbrugge, L. & Sevak, P. Use, type and efficacy of assistance for disability. Journal of Gerontology: Social Sciences, 57B (2002), S366-S379

Compensatory Use of Computers by Disabled Older Adults

Boaz Kahana[1], Eva Kahana[2], Loren Lovegreen[2], and Gul Seçkin[2]

[1] Cleveland State University, Psychology Department, 2121 Euclid Avenue, Cleveland, Ohio 44115-2214, USA
exk@case.edu
[2] Case Western Reserve University, Sociology Department, MTHM 226, 10900 Euclid Avenue, Cleveland, Ohio, 44106-7124, USA

Abstract. This paper focuses on diverse adaptive benefits of computer use by frail elders in the U.S.A. who are attempting to maintain independent lifestyles. Case studies are presented based on data from an ongoing longitudinal study (N=1000) of older adults living in a retirement community [12]. It is argued that technology may extend the autonomy and independence that older people value, which may be threatened as they encounter functional decline, especially in very old age. The data lend support to theoretical models proposed by Baltes and colleagues regarding *selective optimization with compensation*, as characterizing successful adaptation to late life disability. Furthermore, results demonstrate that technology is increasingly utilized by old-old adults and that the flexibility afforded by computer-supported health communication presents a useful resource for older adults.

This paper will focus on diverse adaptive benefits of computer use by frail elders in the U.S.A. who are attempting to maintain independent lifestyles. Case studies are presented from a large panel study (N=1000) of older adults living in a retirement community that does not provide formal services for the elderly. These case studies lend support to theoretical models proposed by Baltes and colleagues regarding *selective optimization with compensation*, as characterizing successful adaptation to late life disability [1].

It has been argued that new technologies offer hope for the elderly to remain independent and to continue engagement in meaningful activities even after they encounter physical and cognitive limitations [2]. New computer technologies and user-friendly gadgets could lead to extending independent living and delaying institutionalization of the elderly [3]. Cell phones, that have become ubiquitous even among older adults, may hold particular promise through use of global positioning systems for assisting elderly persons in finding their way should they get lost due to the onset of cognitive impairment. Sensor-based technology is also being developed to help remind elderly patients to check their blood pressure or take needed medications [4].

Technology use thereby could extend the autonomy and independence that older people value, which is threatened as they encounter functional decline that is brought

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 766 – 769, 2006.

about by the disability cascade, especially in very old age [5]. A critical determinant of acceptance and satisfaction with technological advances relates to the degree of autonomy and privacy that they ultimately provide to the aging individual. Accordingly, there may be concerns with technologically sophisticated monitoring systems that are marketed with the promise of enhancing independence. Older persons may consider these devices to infringe on their autonomy and may decline to use them [6]. In contrast those technological advances that clearly put the older person in the "driver's seat" are likely to be far more rapidly adopted by the elderly.

A particularly important example of consumer directed technology is reflected in the increasing use of email for maintaining communication with significant family members, friends and health-care providers. In our model of successful aging that is described in greater detail in a paper presented at this conference by Eva Kahana, computer use for information gathering and for communication is considered an emergent form of proactive adaptation that contributes to successful aging [7]. We have been particularly impressed by the opportunities afforded through computer use to remain in contact with family members for older adults living in retirement communities at a significant distance from members of their family.

Another important function of computer technology arises as older people face stressful life situations because they are diagnosed with life-threatening illnesses such as cancer [8]. Our research focusing on cancer patients has demonstrated that online discussion and support groups can serve as important resources, particularly for educated older adults that help them cope with both instrumental and emotional challenges posed by a cancer diagnosis [9]. Older patients, who may feel particularly stigmatized by having a life threatening illness, could derive special benefits from going on line. Interestingly, our findings reveal that patients increasingly communicate information they obtain from the Internet to their physicians and reports that physicians are welcoming of such discussions.

In today's increasingly complex health-care environment the ability of health-care consumers to become active partners in their care has become increasingly important. It has been documented that older adults are an underserved group in terms of physician communication, both in terms of preventive services and in care [10].

Patient initiatives and assertiveness in communicating with the doctors can facilitate obtaining better health-care [11]. Specifically older patients who become more assertive in their interactions with physicians may experience better health outcomes. In this regard our research reveals that older age of patient represents a deterrent to cancer screening recommendations by physicians. However those patients who were willing to remind their physicians that they need to be referred for cancer screening were found to receive more recommendations and referrals for cancer screening then their more passive counterparts regardless of age. With extensive media and Internet coverage about the benefits of cancer screening, technologically savvy seniors are more likely to derive health benefits.

Among respondents in our longitudinal study of successful aging, we found notable cohort differences in computer use [12]. Whereas in 1996, only 12% of older adults in our longitudinal study reported using computers, ten years later the proportion of computer users reached 48%. It is important to note that we are reporting cross sectional differences and are not suggesting that the same individuals increased their likelihood of using computers as they aged. Rather, we find evidence that individuals reaching old age at present have been socialized into use of computers.

Data is available on technology use of 123 individuals currently participating in our ongoing longitudinal study of successful aging (mean age=75.6; SD=7.7). This sample is subdivided into an urban component (n=67) and a retirement community-based component (n=56). There appears to be evidence of prevalent technology use but in the context of a digital divide. The use of computers and other technologies such as cell phones are far more prevalent for the more educated and affluent retirement community-based sample than for urban elderly [13]. Thus for example, cell phones were owned by 84% of retirement community-based residents but only 33% of urban dwellers. Those older adults who owned cell phones were reported being frequent users with the majority using their cell phones daily or weekly.

Similar patterns were observed in reported computer use and ownership. Eighty-nine percent of retirement community-based owned a computer with 36% of urban dwellers reporting ownership. About one-fifth of the sample gained access to computers associated with their work. It is notable that even in this old-old population ownership was clearly associated with frequent (almost daily) computer use. Interestingly, most of the elderly report that they are self-taught in the use of computers, thereby illustrating initiative on the part of older adults in adopting technology.

The most frequent goal in use of a computer appears to be that of e-mail and accessing the internet. Seventy-two percent of those using a computer report that they use e-mail and 58% report using the internet. The vast majority of internet users subscribe to an internet service. Urban elderly predominately use the internet for checking on news, weather and sports thereby accessing information of general interest. Older adult living in the retirement community are more likely to use the internet as a resource for obtaining targeted health information and managing their day-to-day lives through use of the internet (e.g., shopping and banking). This data suggests that technology is increasingly utilized by old-old adults. However, those in greatest need of supports may be slower to learn about the potential of computer technology in enhancing their independence than are their more educated and more affluent counterparts.

Considering qualitative comments made by study respondents, computer use was reported to serve a useful function for older adults whose handwriting had deteriorated due to loss of fine motor control. These elderly could maintain communication with friends, family and even physicians via e-mail. Older adults facing newly diagnosed illnesses such as cancer, found it helpful to access information on the internet. Computer use was also reported by older adults with limited mobility to foster connectedness to their informal support network. Moreover, computers also served as helpful tools for health monitoring by older adults and, in some cases, by distant family members and healthcare providers who could be informed about changing health conditions. The most prevalent direct use of computers by frail elders was for obtaining health relevant information. This use allowed for finding health care experts, to learn about drug interactions, and to explore new or alternative therapies for chronic health problems. Respondents reported that their physicians were increasingly willing to reply to patient inquiries via e-mail. For disabled older adults, who may not have adequate access to transportation, the flexibility afforded by computer supported health communication presents a useful resource.

References

1. Baltes, B.B., Dickson, M.W.: Using Life-Span Models in Industrial Organizational Psychology: The Theory of Selective Optimization with Compensation. Applied Developmental Science 5 1 (2001) 51-62
2. McCreadie, C., Tinker, A.: The Acceptability of Assistive Technology to Older People. Ageing & Society 25 1 (2005) 91-110
3. Giraldez, M.C., Casal, C.R.: The Role of Ambient Intelligence in the Social Integration of the Elderly. In: Riva, G., Vatalaro, F., Davide, F., Alcañiz, M. (eds.): Ambient Intelligence: The Evolution of Technology, Communication, and Cognition towards the Future of Human-Computer Interaction. IOS Press Amsterdam Netherlands (2005) 265-280
4. Korhonen, I., Parkka, J., Van Gils, M.: Health Monitoring in the Home of the Future. Engineering in Medicine and Biology Magazine, 22 3 (2003) 66-73
5. Verbrugge, L.M., Reoma, J.M., Gruber-Baldini, A.M.: Short-Term Dynamics of Disability and Well-Being. J of Health and Social Behavior 35 2 (1994) 97-117
6. Kahana, E., Kahana, B., Seçkin, G.: Technology Use and Successful Aging. Paper presented at the Conference on Technology Use and Aging, Univ. of Washington Seattle USA (October 30, 2004)
7. Kahana, E., Kahana, B., Kercher, K.: Emerging Lifestyles and Proactive Options for Successful Aging. Ageing International 28 2 (2003) 155-180
8. Kahana, B., Sterns, S., Kahana, E., Deimling, G.: Comorbidity of Functional Limitations to Perceptions of Stigma among Elderly Cancer Survivors. Presented at the Gerontological Society of America Annual Meeting, Orlando Florida USA (November, 2005)
9. Seçkin, G., Kahana, E.: Patient Participation and Decision Making in the Age of Communication Technology. Paper Presented at the International Conference on Communication in Health Care, Northwestern University, Chicago, Illinois USA (October 2005)
10. Fox, S., Roetzheim, R., Kington, R.: Barriers to Cancer Prevention in the Older Person. Clin Geriatric Med 13 (1997) 79–95
11. Kahana, E. Kahana, B:. Patient Proactivity Enhancing Doctor-Patient-Family Communication in Cancer Prevention and Care among the Aged. Patient Education & Counseling 2075 (2003) 1-7
12. Kahana, E., King, C., Kahana, B., Menne, H., Webster, N., Dan, A., Kercher, K., Bohne, A., Lechner, C.: Successful Aging in the Face of Chronic Disease. In: Wykle, M., Whitehouse, P., Morris, D. (eds.). Successful Aging Through the Life Span Intergenerational Issues in Health. Springer New York New York (2004) 101-126
13. Campbell, R.J.: Older Women and the Internet. J of Women & Aging, 16 1-2 (2004) 161-174

Computer Aids Clients with Psychiatric Disabilities in Cognitive and Vocational Rehabilitation Programs

Tzyh-Chyang Chang[1,2], Jiann-Der Lee[1,*], Shwu-Jiuan Wu[3,4], Ming-Jen Yang[5,6], Chun-Hua Shih[5,6], and Juei-Fen Huang[7]

[1] Department of Electrical Engineering, Chang Gung University, Tao-Yuan, Taiwan
259 Weng Hwa 1st Rd., Kweishan, Taoyuan 333, Taiwan
{tcc200, jdlee}@mail.cgu.edu.tw
[2] Department of Occupational Therapy, Chang Gung University, Tao-Yuan, Taiwan
259 Weng Hwa 1st Rd., Kweishan, Taoyuan 333, Taiwan
[3] Graduate Institute of Medical Informatics, Taipei Medical University, Taipei, Taiwan
250 Wu-Xin Street, Taipei 110, Taiwan
[4] Taipei Veterans General Hospital, Taipei, Taiwan
201, Sec. 2, Shih-Pai Rd., Taipei 112, Taiwan
sjwu123@yahoo.com.tw
[5] Department of Psychiatry, Kaohsiung Medical University, Kaohsiung, Taiwan
100 Shi-Chuan 1st Rd., Kaohsiung 807, Taiwan
[6] Graduate Institute of Behavior Sciences, Kaohsiung Medical University, Taiwan
100 Shi-Chuan 1st Rd., Kaohsiung 807, Taiwan
{m750141, schua}@ksts.seed.net.tw
[7] Department of Foreign Language, Kinmen Institute of Technology, Taiwan
1, Dasyue Rd., Jinning Township, Kinmen County 892, Taiwan
jueifen123@hotmail.com

Abstract. The purposes of this study are to assess the aided effects of computer in cognitive and vocational rehabilitation of clients with psychiatric disabilities and to follow up their employment status. All participants from a community mental rehabilitation unit take a three-month computer skill training program. Participants complete computer key-in test and attention test at the beginning and at the end of the computer skill training program. The researcher assesses all participants' behaviors in class by using observation in every session. After six months, ten participants are still employed and their works are related to computer skills. The significant cognitive improvements of these participants are attention focus ability, problem solving skills, and memory retention ability. In addition, participants completing computer training program can use learned computer skills to obtain more work opportunities. Therefore, applying computer skill training programs to psychiatric disabled clients can improve not only their cognitive abilities but also vocational skills.

1 Introduction

It is to the psychiatric patients' best interest to have prevocational rehabilitation program designed to meet modern job market needs and strengthen their working skills in order

* Corresponding author.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 770 – 776, 2006.

to better enhance their chances of obtaining job offers [1]. There are several job types suitable for psychiatric patients including cleaning service, beauty treatment for automobile, food industry, janitorial, newspaper dispatcher, gardener, production worker, assistant to the secretary, and computer processing related occupations. With the advance of computer in everyday life, not only will it be found in work related environment but also in academic learning, financial systems, national defense, and medical services. Those with knowledge in computer stand a better chance of obtaining job offers and promotions. It is therefore beneficial to psychiatric patients to have computer knowledge and skills in order to obtain higher rate of employment and job diversity [2].

The most common problem faced by in-service psychiatric patients is interference of symptoms that ultimately affect their work performance. In all psychiatric disorders the degeneration of schizophrenia is most eminent. And the general symptoms of schizophrenia can be divided into positive and negative ones. The positive ones include delusion, hallucination, loosening of association, over excitement and exaggeration …etc. The negative ones incorporate affection sluggishness, coldness, socially withdrawal, difficulty in abstract thinking, and non-fluent speech … etc. Generally speaking, psychiatric disorders will interfere with patients' social and occupational functions. Patients with positive symptoms react to the medication better than negative ones. Patients with negative symptoms are more susceptible to lacking of work drive, inability to establish stable intimacy relationships with others rendering inadaptability to family and work life. Many residual symptoms of chronic schizophrenia patients will affect their work performance [3-5].

The major cognitive deficits affecting schizophrenic are attention focus ability, memory retention ability, learning ability and executing ability [6]. In recent years, the focus of psychiatric rehabilitation is the intervention of cognitive rehabilitation. Using individual intervention in clients with cognitive problems improves their cognitive abilities and vocational abilities [6, 7]. In addition, attention plays an important role in work performance. Clients with better attention can relatively meet the meticulous working projects, their work endurance and efficiency will also be better. On the contrary, clients with worse attention are prone to making mistakes at work; their endurance is limited, so is their efficiency. Besides their probabilities of accident happening at work are greater than others [8-10].

If psychiatric disabled clients take the planned prevocational training programs, their job retention rate will be higher and so will their work efficiency [11-13]. Bell [6] uses computers as learning devices in training cognitive functions and work skills of psychiatric disabled clients and gets satisfying results with the training programs. Bell [6] considers that cognitive training courses based on computer skills can improve not only psychosocial functions but also learning behaviors. The best result of vocational rehabilitation will be reached by using computers as assistant devices [6, 14].

Therefore, psychiatric rehabilitation professionals need to plan computer courses in clients' vocational rehabilitation programs in order to assist stable psychiatric patients in increasing their work obtaining rates. Computer skill training programs for psychiatric patients need to be designed to meet modern job market demands and strengthen patients' work abilities. In designing a computer learning course, motivation and course design appropriate to patients' level of competence are extremely emphasized at first. The second point worth considering is how to increase learning outcomes and train their working behaviors at the same time. Finally, clients can apply their learned computer skills in their job and maintain work stability [2, 15].

The purposes of this study are to assess the aided effects of computer in cognitive and vocational rehabilitation for psychiatric disabled clients and to follow up their employment status. Clients with psychiatric disabilities, in this study, participate in computer skill training courses in order to prepare for successful employments.

2 Methods

There are 24 participants in this study. All participants from a community mental rehabilitation unit take the computer skill training program for three months (see table 1).

Table 1. The computer skills training programs in this study

Basic concepts of computer , mouse operation, turn on/off
Basic understanding of windows
File management, English/Chinese key in
The printer install, set up and printing
WordPad learning course
Microsoft Word learning course
Microsoft Excel learning course
PowerPoint learning course
General review

Table 2. The criteria of class performance assessment

SCORE / ITEM	3	2	1	0
Present	On time	late	Leave early	Absent
Emotional disturbance	Stable(0)	Mild(1-2)	Moderate (3-4)	Serious (above5)
Cognition disturbance	Non(0)	Mild(1-2)	Moderate (3-4)	Serious (above5)
Meaningless behavior	Never(0)	Sometimes (1-2)	Usually(3-4)	Always (above5)
Tolerance capacity	Voluntary	≧5	≦5	Avoid
Comprehen-sion	Description 1	Description 2	Description 3	Description ≧3
Responsibility	On time	Above≧ 50%	≦50%	Fail
Sustention	3hours	2hours	1hour	<1hour
Obey	obey	pretend	cold	reject

Participants completed computer key-in test and attention test at the beginning and at the end of the computer skill training program. The researcher assesses all participants' behaviors in class by using observation – class performance assessment (see table 2) in every session. The class performance assessment is based on comprehensive occupational therapy evaluation scale (COTES) [16,17].

CAI key-in software is used in both the first and the last session of this study to assess key-in skill. The software can record accuracy and total number of word counts for English or Chinese typing. In addition, Chu's attention test is used in this study as well to assess participants' attention [17-19].

3 Results

Participants' information in this study is shown in table 3. In addition, the difference of pre/post test for subjects' computer skill training is shown in table 4.

Table 3. The information of subjects（n=24）

Age	M=31.75	SD=5.86
Sexual		
Male	N=10	41.7%
Female	N=14	58.3%
Marry status		
Married	N=1	4.2%
Single	N=23	95.8%
Work experience		
Non	N=20	83.3%
Yes	N=4	16.7%
Computer-learning		
Non	N=14	58.3%
Yes	N=10	41.7%
Diagnosis		
Schizophrenia	N=18	75%
Emotional disturbance	N=6	25%
Educational level		
Junior high school	N=2	8.4%
Senior high school	N=11	45.8%
College	N=11	45.8%

After descriptive statistical method data analysis and comparison of pre/post training assessment, we find that 79.17% (N=19) of participants gain noticeable improvement on computer key-in skill (both accuracy and total number of word counts for English and Chinese typing increase).

With more in-depth analysis, we find that 95.83% of participants' Chinese typing word counts steps up (two words for the least improved, 206 words for the most improved, M=80.54, SD=66.24); 95.83% of the participants' Chinese typing accuracy raises from 1% to 5% (M=1.88%, SD=1.73%) variously; 95.83% of the participants' English typing word counts augments (fifteen words for the least improved and 221 words for the most improved, M=115.63, SD=66.92); 83.33% of the participants' English typing accuracy grows from 1% to 17% variously (M=1.54%, SD=6.33%). (see table 5).

Table 4. Difference of pre/post test for subjects' computer skills training（N=24）

Content	Paired t-test	P value
Attention	-2.246	.035*
Chinese key-in	-5.957	.000**
Accuracy of Chinese key-in	-5.318	.000**
English key-in	-8.464	.000**
Accuracy of English key-in	-1.193	.245**

*p<0.02, **p<0.01

Table 5. The change of subjects' computer key-in skills（N=24）

Increasing item	Increasing percentage	Mean	SD
Chinese key-in	96.10%(N=23)	80.54	66.24
Accuracy of Chinese key-in	96.10%(N=23)	1.88%	1.73%
English key-in	96.10%(N=23)	115.63	66.92
Accuracy of English key-in	83.10%(N=20)	1.54%	6.33%

The results of this study have 66.67% (N=16) of participants get improved in their attention ability and 79.17% (N=19) of participants perform improved key-in skills. The researcher assesses all participants' behaviors in class by using observation in every session. The three most significant improvements of participants' behaviors are comprehension (N=21), frustration tolerance (N=19), and endurance (N=17).

4 Discussion

We find that both attention focus ability and computer key-in skill are improved enormously for one married participant after receiving the training program; four participants with work experience previously gain improvement on computer key-in skill; and nine out of ten participants who have prior knowledge about computer also improve a lot after the training. However, participants with no work experience or prior knowledge about computer perform fairly well on key-in skill after the training, rendering inability to find crucial factor affecting computer key-in skill performance. Besides, there's no prominent relation observed in the study between good key-in performance and attention focus improvement regardless of participants' gender, education background, marital status, age, and diagnosis.

Ten participants are employed and lots of their work duties are related to computer skills. Therefore, applying computer skills training to psychiatric disabled clients can increase not only computer skills but also job choices. After six month's training, ten participants are still employed and their works are related to computer skills (see table 6).

Table 6. Work status after completing computer-training course (N=10)

Job type	N	Description
Computer-related job	7	computer assistant, medical administrator, assistants, bus center
Cleaning company	2	Worker assistant
Cable TV company	1	Self-working

The significant cognitive improvements of these participants are attention focus, problem solving skills, and memory retention. In addition, participants completing the computer training program can use learned computer skills to obtain more work opportunities.

Psychiatric disabled clients receiving computer skill training stand a better chance of obtaining works and work diversities. Furthermore, during follow-up visits to the twenty-four subjects receiving the training, we find that the four most significant improvements psychiatric disabled clients experienced are (1) work endurance improvement (37.5%, N=9); (2) self-affirmation improvement (33.3%, N=8); (3) problem-solving ability improvement (25%, N=6); (4) reaction ability improvement (20.8%, N=5).

The limitation of this study is addressed as follows. Since only one community mental rehabilitation unit is studied, and the sampling number is quite limited, and there are no other controlled groups to compare to, the result of this study can not be generalized to all psychiatric patients' computer-related skill training application.

However, the study results provide clinical psychiatric rehabilitation professionals with diversity for inclusion of computer skill training with patients in their rehabilitation programs in the future.

5 Conclusion

After receiving the training, not only their attention and computer skills are improved also their comprehension abilities, frustration tolerance, and endurance are enhanced tremendously. The conclusion of this study can provide medical rehabilitation professionals with some ideas which can help incorporate computer learning courses for stable patients in vocational rehabilitation processes. It can provide clients with specialized training opportunities to strengthen their capabilities of obtaining works in the future as well as to increase their rate of returning to society in a shorter duration.

References

1. Stauffer, D.L.: Predicting Successful Employment in the Community for People with a History of Chronic Mental Illness. Occupational Therapy in Mental Health, 6(2). (1986) 31-49
2. Lau, A.: Computer Applications in Occupational Therapy: A Journal Review. In Cromwell, F. S. (ed.), Computer Applications in Occupational Therapy (pp. 93-98). New York: The Haworth Press (1986)
3. Liberman, R.P. (ed.): Psychiatric Rehabilitation of Chronic Mental Patients. Washington. D. C.: American Psychiatric Press (1988)
4. Torrey, E.F.: Surviving Schizophrenia: A Manual for Families, Consumers, and Providers (3rd. ed.). New York: Harper Perennial (1995)
5. Durham, T.: Work-Related Activity for People with Long-Term Schizophrenia: A Review of the Literature. British Journal of Occupational Therapy, 60(6). (1997) 248-252
6. Bell, M., Bryson, G., Greig, T., Corcoran, C., Wexler, B.E.: Neurocognitive Enhancement Therapy with Work Therapy. Arch Gen Psychiatry, 58 (2001) 763-768
7. Wykes, T., Reeder, C., Corner, J., Williams, C., Everitt, B.: The Effects of Neurocognitive Remediation on Executive Processing in Patients with Schizophrenia. Schizophrenia Bulletin, 25 (1999) 291-307
8. Nuechterlein, K.H., Dawson, M.C.: Information Processing and Attentional Functioning in the Developmental Course of Schizophrenic Disorders. Schizophrenia Bulletin, 10(2) (1984) 160-203
9. Sereno, A.B., Holzman, P.S.: Spatial Selective Attention in Schizophrenia, Affective Disorder, and Normal Subjects. Schizophrenia Research, 20 (1995) 33-50
10. Medalia, A., Aluma, M., Tryon, W., Merriam, A.E.: Effectiveness of Attention Training in Schizophrenia. Schizophrenia Bulletin, 24(1) (1998) 147-152
11. Becker, D.R., Drake, R.E., Farabaugh, A., Bond, G.R.: Job Preferences of Clients with Severe Psychiatric Disorders Participating in Supported Employment Programs. Psychiatric Services, 47(11) (1996) 1223-1226
12. Blankertz, L., Robinson, S.: Adding a Vocational Focus to Mental Health Rehabilitation. Psychiatric Services, 47(11) (1996) 1216-1222
13. Siu, A.H.: Predicting Employment Outcomes for People with Chronic Psychiatric Illness. Occupational Therapy in Mental Health, 13 (1997) 45-58
14. Galletly, C., Clark, C.R., MacFarlane, A.C.: Treating Cognitive Dysfunction in Patients with Schizophrenia. Journal of Psychiatry & Neurosci, 25 (2000) 117-124
15. Shaw, C. McKenna, K.: Microcomputer Activities for Attention Training. Cognitive Rehabilitation, 7(1) (1989) 18-21
16. Brayman, S.J. et al.: Comprehensive Occupational Therapy Evaluation Scale. American Journal of Occupational Therapy, 34 (1976) 328-331
17. Hsiao, S.J., Pan, A.W., Chung, L., Lu, S.j.: The Use of Evaluation Tools in Mental Health Occupational Therapy: A National Survey. Journal of Occupational Therapy Association, R.O.C., 18 (2000) 19-32
18. Chu, T.F.: The Application of Vocational Assessments Tools in Psychiatric Patients. Journal of Occupational Therapy Association, R.O.C., 9 (1991) 43-61
19. Chu, T.F.: The Manual of Chu's Attention Test (1997)

Ethically Aware Design of a Location System for People with Dementia

Roberto Casas[1], Álvaro Marco[1], Jorge L. Falcó[1], José I. Artigas[1], and Julio Abascal[2]

[1] TecnoDiscap Group,
University of Zaragoza
50018 Zaragoza, Spain
{rcasas, amarco, jfalco, jiartigas}@unizar.es
http://diec.unizar.es/~gtd/
[2] Laboratory of HCI for Special Needs,
University of the Basque Country – Euskal Herriko Unibertsitatea
20018 Donostia, Spain
julio.abascal@ehu.es
http://www.sc.ehu.es/acwbbpke/

Abstract. This paper presents an experience of ethically aware design of a location service intended to monitor residents in an institution for people with dementia. As location systems have a direct impact over privacy and personal autonomy, the system design was preceded by a deep study of ethical and social impact. Among its contributions it can be emphasized that this study was extended both to patients and also to caregivers that were located in order to provide urgent assistance. The conclusions of this experience were compiled as design guidelines and used for the technological design.

1 Introduction

Assistive Technology has frequently a large impact over ethical and social issues. Designers usually concentrate on technological aspects and tend to ignore the impact of their designs. It is claimed that social and ethical issues are not under the responsability of scientific staff and should be taken into account by policy makers, caretakers, social workers, etc. Nevertheless, it is well known that the technical design of many devices and services include characteristics affecting the rights of the users that can not be removed because are substantially rooted in the conception of the application.

In order to avoid eventual negative impact over the user rights caused by Assistive Technology, the design process should include the following conditions:

Designers should

– be aware of the possibility of producing technology having side effects over social and personal issues.
– know how to identify and face ethical issues, or be able to contact experts that can do this task.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 777–784, 2006.

– have at their disposal design methodologies and tools that do no impede ethical aware design.
– master methodologies and tools for evaluation and test of the results including evaluation of ethical impact.

Let us deeper describe these conditions.

1.1 Awareness of Technology Ethical Side Effects

The first condition to avoid ethical impact is that designers are aware of the possibility of creating products that have undesired negative effects over the privacy, autonomy, security, etc., of the users. Technologists have to know the importance of the ethical issues and to understand that most frequently side ethical effects are involuntarily included in the design for technological reasons. In addition, they must know that when these issues are part of the design, it is almost impossible to remove them after the completion of the process.

Ethically aware design guidelines are very helpful to assist designers with short experience in this field [1].

1.2 Identification and Solution of Ethical Issues

The analysis of the potential ethical impact requires social expertise that may be obtained from external experts and sometimes from caregivers, users and relatives. Participation of multidisciplinary teams can be required in order to combine technological and sociological backgrounds. The exhaustive analysis of application scenarios can help to understand the future effect of the designed technology. In this way a sorted list of potential ethical impacts can be produced. A detailed description would be useful in order to help designers to avoid the detected dangers.

1.3 Design Methodologies and Tools Ethically Respectful

It frequently happens that the own methodology used for the design (or the tool that gives support to it) forces to decisions that are ethically doubtful. For instance, the type and quantity of data gathered and stored, the way they are obtained and processed, the possibility of controlling who has access to them, etc., may be conditioned by the tool or the platform adopted by the development team, and can make impossible to establish controls to avoid ethical side effects. Therefore, it is necessary to develop methodologies and tools to support them, which are flexible in how to process sensible data, and do not impose a specific approach.

1.4 Methodologies and Tools for Evaluation and Test of the Ethical Impact

When the design is completed, a thorough test of the possible ethical impact should be carried out. At this stage, evaluation guidelines are required to help the evaluator in the analysis. Large applications may require tools that support the evaluation, collection of the results and redaction of the report.

Ethical impact evaluation may sometimes be combined with usability and accessibility evaluations.

2 Privacy Issues in Location Systems

Location services have a characteristic impact over the user rights that affects not only to people with disability, but any person carrying a locatable device, such as a mobile phone, for instance.

2.1 Considerations for General Location Services

Focusing on people location systems, the most evident ethical impacts are related to privacy and autonomy. Apart from legal aspects, it is also necessary to consider that the fact to know the location of a person in a precise way may have a very important psychological effect; the users can associate it with the "big brother" phenomenon [2]. This feeling usually decreases when the users receive an added service or when they know the advantages in security through being temporary or permanently located. On the other side, caregivers or other workers —that could be located to enhance their service— normally reject to be permanently located. It is evident that acceptation is related to the benefits that the person expect from the location system. Even that, acceptation by patients can be in some cases forced because rejection would means the loss of other services.

There is a debate due to the relatively recent emergence of commercial location systems. This discussion has been intensified by the EU [4] and US [5] Directives that request the location of the mobile terminals calling to the emergency services. The European Directives 95/46/EC and 2002/58/EC [3,4] distinguishes location and content data, but some authors, such as Escudero [6] opine that they have the same sensitivity. Using simple analysis techniques from location data, it is easy to determine where, when, how long and with whom the people have been. In other words, obtain behavioural patterns, consumption habits, etc. that is to say, personal information.

These directives states that location data may only be processed when there is an explicit consent by the located person, when the data harvesting is anonymous (it is not possible to relate the person and its location) or when the service is applied by organizations dealing with emergency calls. This last point coincides with the Federal Communications Commission Directive of the United States requesting that the mobile terminals calling to the emergency service 911 can be located [5].

An important contribution comes from Location working group of Open Mobile Alliance (OMA) (former Location Interoperability Forum), the leading industry forum for developing market driven, interoperable mobile service enablers. Considering the indications of the European directives in terms of data protection, they propose recommendations for a location services designer to rely on. It is absolutely necessary that the harvesting of the location information is known and understood by all the users of the system, who must give explicit consent. In the same way, the use that the information is going to have and the time that is going to be kept must be clear. Obviously, the security has to be adapted to the importance and sensitivity of the data. They also indicate that when the location is a legal requirement, privacy considerations can be avoided [7].

In order to guarantee the location privacy, security in information storage medias and communication channel —mainly wireless— are subjects of great practical importance. Beresford and Stajano [8] review the state of the art on this subject and suggest diverse directions of future investigation. They propose to increase security by using pseudonyms in such a way that only the user knows her/his own identity. It is also proposed to increase the entropy of the information to make the valid data extraction more difficult, for example with the use of many fictitious users.Myles et al. [9] suggest another approach through systems that offer the location data based on configurable parameters: the identification of the user who demands the information, the hour of the day, location and preferences of the user, etc. They propose a middleware level that operates in between the positioning system and the applications using the location service, validating their requests.

2.2 Specific Considerations for Assistive Technology

Pompano [10], shows that the designed laws to protect the people with disabilities, often reduce their privacy. The use of a positioning system entails an obvious utility for the people using it, but due to their special characteristics, a reflection becomes necessary on the ethical aspects related to the invasion of the private life and the restriction of freedom that it could suppose. Even more, if we thought that in many cases, although the informed consent could be signed, we realize that the person does not completely understand the privacy implications of location.

As in other technology fields (voice recorders, camcorders, etc.), the problem is not in the positioning system, but in the misuse of the information it offers. In the same way that the location of mobile phones is not the same if the information is used with commercial aims or to locate an emergency situation, within the assistive technology, there is a great dependence on the way to manage each situation depending on the application and personal circumstances.

Taking into account the capacities of the person, we can distinguish two different assumptions:

- If the person is able to understand the situation and show her/his consent, the own person accepts or rejects the service. In this case the ethical requirements refer to the treatment given to the data issued by the system and the validity of the way of informing the user to ask his/her consent. It would be necessary to establish a protocol detailing the information to be gathered, the use planned for it, how and how long it is going to be stored or destroyed, who will have access to it, etc. This protocol would have to be explained to the user clearly in order to have his/her informed consent.

– When the person is not able to declare his/her consent, a conflict between two rights is established: in the one hand the right to the intimacy and by another one the right to the health and even to the personal well-being [10]. This could be solved, like it happens in the medical ambit, prioritizing the protection of the person over the privacy right. Although to the main objecvtive is the protection of the person, it would be necessary that the one holding the legal guardianship on the user with disability gives his/her consent. In the OMA guidelines an analogue problem is stated: the tracking of the under aged children by their parents. They indicate that only the informed consent of legal tutor is necessary, but recommend at least the notification to the minor [7].

3 Case Study

Systems used to locate individuals have a high impact on rights to privacy, including the invasion of private life and the restriction of personal freedom. Among other conditions, awareness and explicit consent are claimed to be the main ways to protect users from privacy invasive systems [11], but when users are people with cognitive disabilities, consent and awareness are frequently not possible.

This section presents a study conducted within the Dalma project to detect ethical problems and to generate criteria for an ethically aware technical design.

Dalma project developed an alarm and event notification system with location capacity, for use in residences, hospitals and other public or private institutions [12]. The designed location system achieves a precision of a few centimetres based on small wearable devices using radio-frequency and ultrasonic technologies. This system was enhanced with the sensing of body parameters, and the integration into a Smart Home in two separate projects, Heterorred I and II [13].

The main benefit of Dalma is to provide safer and greater mobility to the person by detecting risk situations, such as staying at a dangerous place for too long time (leading to a risk of falling), wandering (the possibility of disorientation), absence or lost contact (missing from the residence), repetitive actions (possible behaviour denoting anxiety, escapism or other pathologies), etc. Indirect benefits are the extension of surroundings that are considered safe for the person, psychological reinforcement, and the reduction of time waiting for assistance.

3.1 First Technical Design

Rooted in previous experiences [14,15], a first design of the Dalma system took into account only technological issues. In this way, an indoors location system using radio-frequency and ultrasonic technologies was specified, which was able to detect the location of a small device worn by each user.

The developed positioning system consist of several cells covering location areas, each one composed by several fixed devices denoted *beacons*, and several mobile devices denoted *tags*. To perform localization, the beacons successively send ultrasonic pulses, and the tags receive them and measure the time of flight of these chirps. After that, the times of flight stored by each tag are collected via radio-frequency (Bluetooth) and sent to a central control unit (PC), where the position of each tag is calculated. The system was able to locate typically around 20 tags in cells of 100 m^2, with an accuracy of 5-10 cm, and a temporal resolution of one new location every second. These features are enough to the alarm system requirements, allowing location and tracking of the users with good reliability.

3.2 Ethical Design Guidelines

Ubiquitous technology allows logging many human parameters such as location, movements, communication tasks, body parameters, etc., without the consent or even the awareness of the observed person. As we have stated in section 2, this situation requires a suitable legal frame, but relevant legislation is different, and sometimes insufficient, in each country. There are several studies [2,7,8,9,16] about privacy issues for location systems, but they are not always applicable when users have cognitive disability. In this particular case (and also frequently in the development of other Assistive Technologies) Pompano [10] argues that the laws issued to protect the people with disabilities often reduce their privacy.

The use of a positioning system provides an obvious utility for the located person, but due to their special characteristics and dependencies, it is even more critical to consider the ethical aspects related to the invasion of private life and the restriction of freedom. General guidelines cannot be applied because, for instance, many users cannot provide their consent. And even when the informed consent can be signed, it is possible that the user does not completely understand their rights to privacy and the implications of location devices.

Dalma designers understood the potential and danger of this technology and decided to develop internal ethical design guidelines to help themselves to drive technology into a responsible and socially acceptable direction. Following the general privacy protection guidelines stated in section 2, we agreed the following guidelines for technical design criteria:

Consent for location: Signature of explicit consent is an unavoidable requirement that should be fulfilled. As the users will have different cognitive capacities, we distinguish different user profiles:

- Caretakers, as employees, could be positioned provided a signed explicit consent [7]. However, in order to reduce the stress caused by the “big brother” effect, they are only located when an alarm occurs.
- People having sporadic crisis that can be detected by the system or people with accident risks (elderly) are only located when the system detects an alarm situation (fall, disorientation...) or when they request help voluntarily. In this case the positioning information is not stored; it is used solely to deal with the situations effectively.
- Finally we have the people who are periodically positioned in order to detect risk situations (repetitive itineraries, stances in dangerous places...). Usually this people will not be able to give an informed consent, so their

legal tutor should sign it, but always paying attention to protect them. In this case, the location information could be stored the time necessary to be analyzed and then it should be destroyed.

Storing: The location information of caretakers and people having sporadic crises should not be stored indefinitely; instead, it may be used solely to deal effectively with the risky situation at hand. Location information relevant to people who were continuously being located could be stored only for the time necessary to be analysed and then it should be destroyed.

Data protection: As the location system is centralised, database encryption and security access controlled by passwords must be used.

3.3 Redesign of the Proposed System

The implemented version of the Dalma system has all the requested location functions. In addition, privacy and security matters have been enhanced: the inclusion of database encryption and passwords for centralised information; use of pseudonyms to keep the anonymity of the located people; use of secure wireless protocols (implemented in Bluetooth specification) and 128 bits data encryption for wired data through Ethernet, to avoid personal information data hacking. Location monitoring was limited to the following situations: detected risk by alarm systems, or requested by the individual user. User profiles, specifying a different treatment of location information for each type of profile, were defined for caretakers, people with accident risks, and people with cognitive disabilities.

The Dalma system has been deployed in the institution Virgen del Pueyo in Zaragoza (Spain). Several forms were designed to identify and collect technical and ethical issues, but no ethical issues were raised during the evaluation phase of the technology.

4 Conclusions

Devices and services designed for people with cognitive disabilities frequently have impacts over the privacy and autonomy of the users. These effects are deeply rooted in the technological design and can hardly be removed after the development of the system. Nevertheless, technologists tend to ignore these issues, due frequently to the lack of knowledge and supporting methodologies, guidelines, tools, etc.

This paper shows an experience to demonstrate that interested technologists can face this problem analyzing the possible impact of the system, collecting guidelines to assist the design and adopting a ethically aware development approach.

Acknowledgments

This work was supported in part by the Spanish MCYT under Project FIT-150200-2000-215 and by DGA under project T44.

References

1. Abascal J., Nicolle C.: Moving towards inclusive design guidelines for socially and ethically aware HCI, Interacting with Computers. Vol. 17, Issue 5, September 2005. pp. 484–505.
2. Kaasinen, E.: User needs for location-aware mobile services, Personal and Ubiquitous Computing, vol. 7 n^o 1, 2003, pp. 70–79.
3. European Commission, Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, 1995.
4. European Commission, Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications), 2002.
5. Docket, CC, No.96-264, Revision of the Commission Rule to ensure compatibility with enhanced 911 emergency calling system, Federal Communications Commission Reports and Orders, United States, 1996.
6. Escudero, A.: Location data is as sensitive as content data, EU Forum on cybercrime, Brussels, 2001.
7. Location working group of Open Mobile Alliance (OMA) (formerly Location Interoperability Forum), Privacy Guidelines, URL: `http://www.openmobilealliance.org/`, 2002.
8. Beresford, A., Stajano, F.: Location Privacy in Pervasive Computing, IEEE Pervasive Computing, vol. 2, n^o 1, 2003, pp. 46–55.
9. Myles, G., Friday, A., Davies, N.: Preserving Privacy in Environments with Location-Based Applications, IEEE Pervasive Computing, vol. 2, n^o 1, 2003, pp. 56–64.
10. Pompano, J.R.: Privacy Issues and Disabled Persons, Yale-New Haven Teachers Institute, 2000.
11. Langheinrich, M.: Privacy by Design - Principles of Privacy Aware Ubiquitous Systems, International conference on Ubiquitous Computing, 2001, pp. 273–291.
12. Casas, R., Marco, A., Falcó, J.L., Gracia, H. and Artigas, J.I.: DALMA. Location Aware Alarm System for People with Disabilities, 10th Int. Conf. on Computers Helping People with Special Needs, 2006.
13. Sevillano, J. L., Falcó, J., Abascal, J., Civil, A., Jiménez, G., Vicente, S., Casas, R.: On the Design of Ambient Intelligent Systems in the Context of Assistive Technologies, In Miesenberger K., Klaus J., Zagler W., Burger D. (eds.): Computers Helping People with Special Needs. LNCS 3118. Springer, 2004, pp. 914–921.
14. Falcó, J.: Adapted Environments for Elderly People and People with Disabilities: Location and Alarms for people with Dementia, (In Spanish). Doctoral Thesis Report. University of Zaragoza. 1997.
15. Casas, R.: An Interoperable System for Indoors Location Applied to Assistive Technology, (In Spanish). Doctoral Thesis Report. University of Zaragoza. 2004.
16. Clarke, R.: Person-Location and Person-Tracking: Technologies, Risks and Policy Implications, International Conference on Privacy and Personal Data Protection, 1999, pp. 131–150.

TouchStory: Towards an Interactive Learning Environment for Helping Children with Autism to Understand Narrative

Megan Davis[1,*], Kerstin Dautenhahn[1], Chrystopher Nehaniv[1], and Stuart D. Powell[2]

[1] Adaptive Systems Research Group, School of Computer Science
M.Davis@herts.ac.uk
[2] Department of Education
University of Hertfordshire
Hatfield, Hertfordshire, AL10 9AB, UK

Abstract. Children with autism exhibit a deficit in the comprehension and creation of narrative which impacts their social world. Our ongoing research agenda is to find ways of developing interactive learning environments which enhance the ability of individual children with autism to deal with narrative and thus the social world. The study reported here involved 12 children in a longitudinal study which focussed on identifying the particular *aspects of narrative* which individual children found difficult. Our aim was to investigate each individual child's understanding of 'primitive' components of narrative by means of an interactive software game called TouchStory which we developed for this purpose. Our results show, for most of the children, an ongoing and clear distinction in their understanding of the various narrative components, which relates their narrative comprehension as shown by a picture-story based narrative comprehension task.

1 Introduction

Autism is a lifelong pervasive developmental disorder affecting social ability; people with autism exhibit impaired social interaction and communication, and have a limited range of imaginative activities, collectively referred to as the *triad of impairments* [1-3]. In addition, research has shown that children with autism exhibit a deficit in the comprehension and creation of narrative; particularly in understanding the motives and emotions of characters, and the reasons why events happen [4-6]. It is argued that this narrative deficit underlies the social difficulties found in autism [7-9]. Our ongoing research agenda is to find ways of developing interactive learning environments which enhance the ability of individual children with autism to deal with narrative and thus the social world. The study reported here involved 12 children in a longitudinal study which focussed on identifying the particular *aspects of narrative* which individual children found difficult. Our aim was to investigate each individual child's understanding of the 'primitive' components of narrative by means of an interactive software game called TouchStory which we developed for this purpose. The work

* Corresponding author.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 785 – 792, 2006.

falls under the umbrella of the Aurora project [10] which investigates the potential enhancement of the everyday lives of children with autism through the use of robots and other interactive systems as therapeutic or educational ‘toys’.

2 The Longitudinal Study: Concepts and Methodology

Our overall aim is to enhance the abilities of children with autism to understand straightforward narrative. That is, the recounting of a sequence of events involving purposeful characters, such that there is some break in the scenario which makes the story worth telling [11]. Our aim in this study was to investigate children’s understanding of *primitive components* of narrative. We identified these as the characters, the settings or background, and the sequence of events; further, we argue that a narrative sequence is a special form of temporal sequence therefore we include both temporal sequences and sequences with no temporal dimension which we call reversible sequences. This gives narrative and proto-narrative types as follows:

- *type c* which addresses character variability and continuity;
- *type b* which addresses background or setting variability and continuity;
- *type rs* which addresses the sequencing aspects of narrative in the simplified form of a reversible sequence;
- *type ts* which addresses temporal aspects of narrative;
- *type ns* which are complete mini-narratives.

We are concerned with the children’s understanding of narrative *per se*, we are not concerned with literacy; therefore we work with picture narratives. We introduce the term *t-story* to mean a picture narrative or proto-narrative presented by TouchStory. T-stories were prepared for each primitive type, an example of the type *rs* is shown in Fig. 1. The study used a set of 56 t-stories moderated for correctness and lack of ambiguity by a panel of 10 adults. The panel consisted of 7 men and 3 women, with a range of technical experience in using computers; they had no previous involvement with the project or knowledge of the children involved.

The longitudinal study took place in a day school unit for children with impaired communication. All 12 children of the unit were involved, 10 of whom, all boys, were diagnosed either with autism, or behaviours suggestive of autism. We do not claim that these children are representative of all children with autism, but consider them as individual cases, from which some generalisations may be made. Of the remaining children, both girls, one diagnosis was ‘general developmental delay’ and the other ‘social interaction difficulties’. The children were aged between 5 and 11 years.

Twelve visits were made to the unit between February and June 2005; of these one visit was devoted to a paper-based picture-story comprehension task based on the work of Paris & Paris [12], and one to ensuring the children realized the study was at an end. TouchStory was used on the other visits. The longitudinal study uses an *adaptive phase* where the stories presented by the system varied depending on the interaction history of the child, i.e. the scores that child had achieved on previous visits. The goal of this adaptation was to tailor the system towards each child’s individual learning needs, while still providing and enjoyable and rewarding experience. More details on the overall motivation for the adaptation and initial observations are described by Davis *et al* in [13].

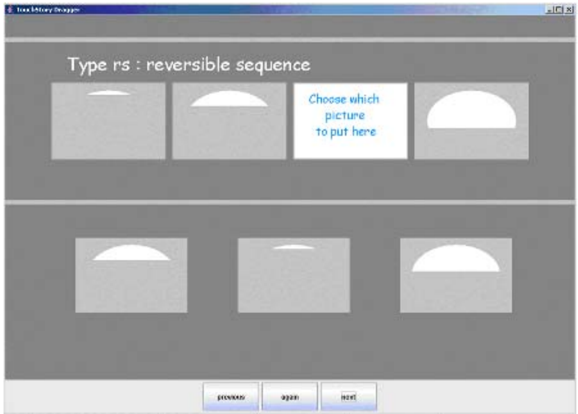


Fig. 1. TouchStory displaying a t-story from the category 'reversible sequence'

Children with autism typically have difficulty with generalising from one context to another, and with remote object references (both relevant to using a mouse), and so TouchStory was presented using a touch sensitive screen. Additionally the children prefer highly predictable environments, therefore new t-stories and new t-story types were introduced gradually as the study progressed. A typical session would consist of 12 to 14 t-stories. For each t-story the task is to select from the bottom row (see Fig.1) the most appropriate picture to complete the top row, and drag it into place. The reward for a correct answer is simply that the bottom row is removed, leaving just the completed sequence. Thus the reward re-enforces the correct answer and does not introduce extraneous sounds or effects; this we consider important in the case of software designed for children with autism as they are particularly likely to focus on seemingly irrelevant details. TouchStory allowed the children controlled autonomy during a session; for example children could use the 'next' button to move on to the next t-story, but repeated clicking of 'next' had no effect, in this way seemingly meaningless repetition, commonly observed in children with autism, was avoided.

Our approach was to incorporate TouchStory sessions into the child's school routine, in a natural, child-centred way. The emphasis throughout the study was on a playful and enjoyable context. The children's communication therapist was present during sessions and gave help or autonomy as appropriate to the individual child. Her advice on whether it was appropriate to terminate a child's session early was followed without question. This approach ensures that TouchStory sessions integrate well with the children's previous experience, that the children's immediate and changing needs are addressed, and that results obtained are relevant to later general classroom-based use of TouchStory.

3 Results

The results presented here focus on the children's *individual* abilities within the 'primitive' narrative categories. The number of t-stories seen differs from child to child because of: the adaptive nature of the software; early termination of some sessions for some children; and absence from school, and so analysis was carried out on the *proportion* of t-stories answered correctly in each category. The median number of t-stories seen by a child (including repeated t-stories) was 100, and the median number of distinct t-stories seen was 45.

A cluster analysis, using single linkage Euclidian distance, of the proportion of t-stories answered correctly in each category, during the study as a whole, is presented in Fig2.

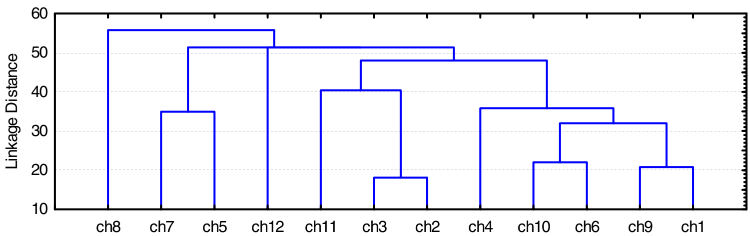


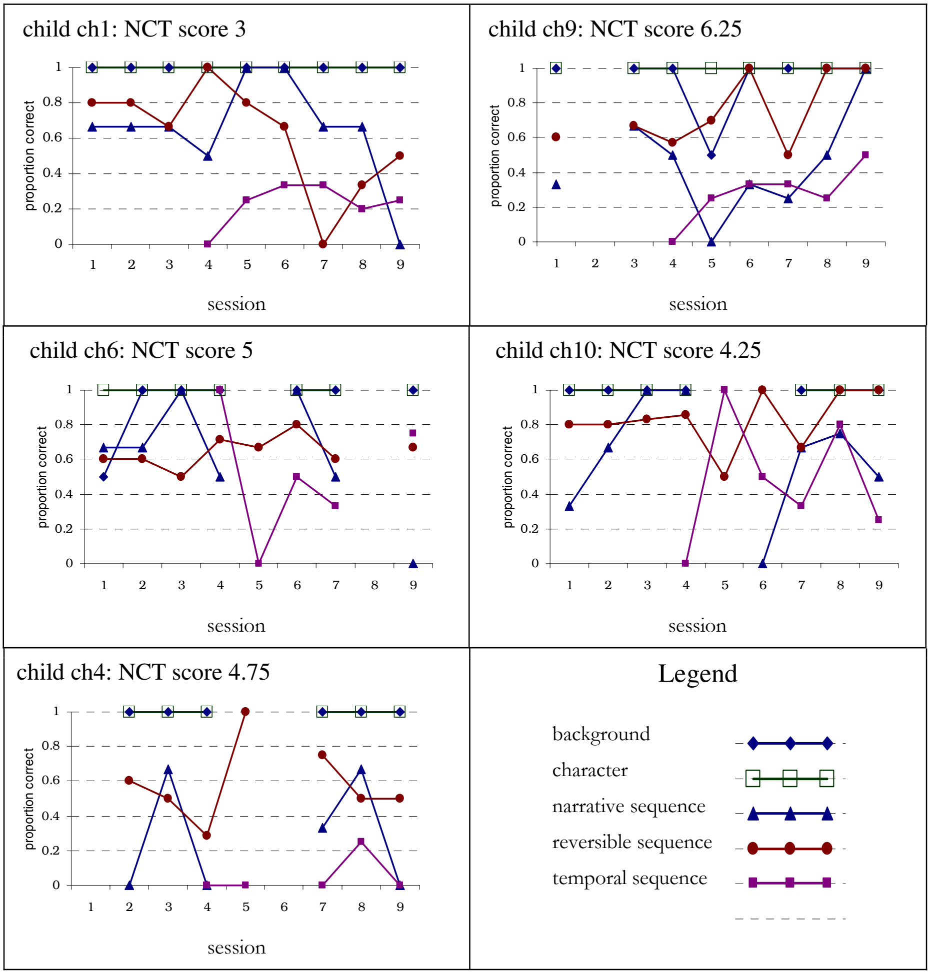
Fig. 2. Cluster analysis of the children's scores in each narrative category

This cluster analysis is used to structure the presentation of the children's understanding of the narrative components presented by TouchStory. To ground the results in activities relevant to everyday school life, results are also presented from the paper-based picture-story comprehension task (NCT), this was scored by two raters independently and the inter-rater reliability was found to be high. The maximum score available for the NCT task was 20; the score for each child is presented as a number out of 20.

The first cluster presented is the {ch1, ch9, ch6, ch10, ch4} cluster. The children in this cluster do well with background and character t-stories and least well with the *temporal sequence* category. All the children in this cluster were diagnosed with autism or behaviors suggestive of autism. Table 1 presents a graph for each child in this cluster showing, for each session which the child attended, the proportion of t-stories answered correctly in each category. The graphs show a clear distinction in performance in different categories, and, in comparison with cluster 2, low variance within categories.

The second cluster presented is the {ch2, ch3, ch11} cluster. Although these children were at their best with the background and character t-stories they did not have the high scores typical of children in cluster1, also, in contrast to cluster 1, they did least well with the *narrative* category. Two of these children do not have diagnoses of autism; the third, at the time of the trial was diagnosed with autism, but this has recently been questioned. He certainly has other issues including dyspraxia and a short attention span. The graphs for these children can be found in Table 2. They show much greater variance within each category than the first cluster and less differentiation among categories.

Table 1. For each session attended the proportion of t-stories answered correctly in each category by each child in cluster 1



The third cluster presented is the {ch5, ch7} cluster. These children were competent in all categories. As these children generally got the answers right, the graphs show low differentiation between categories, and low variance within them; see Table 3. Both children had a diagnosis of autism; however child ch7 seems to have a sound grasp of narrative, he scored highly on the NCT task, he enjoys stories and his teachers find him an empathetic child. He left the unit to rejoin mainstream schooling during the study. Child ch5 did not find TouchStory trivial, he took extreme care over his answers; getting the right answer was very important to him.

Lastly Table 4 shows graphs for the two children with results which do not cluster closely with any other. These two children are special cases; child 12 was generally unhappy throughout the study and was not interested in TouchStory, although in a previous prototype study he had been engaged and successful. Child 8 often found it difficult to focus, additionally although he was able to drag pictures across the screen he was not able to place them as the task required. In his case the first picture he moved in a definite manner was taken as his intended answer.

Table 2. For each session attended the proprotion of t-stories answered correctly in each category by each child in cluster 2

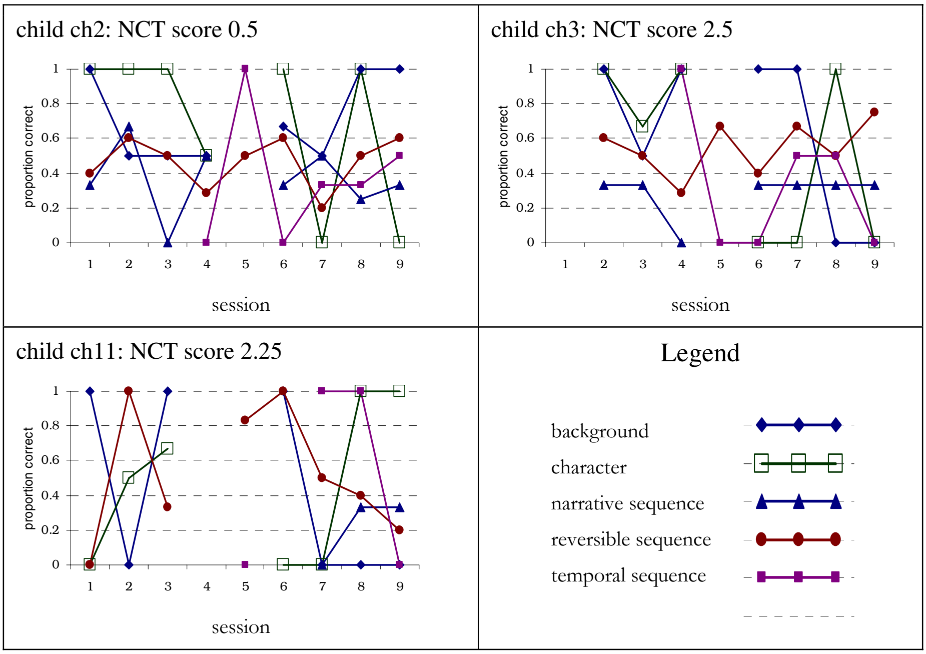


Table 3. For each session attended the proportion of t-stories answered correctly in each category by each child in cluster 3, for legend please see Table 2

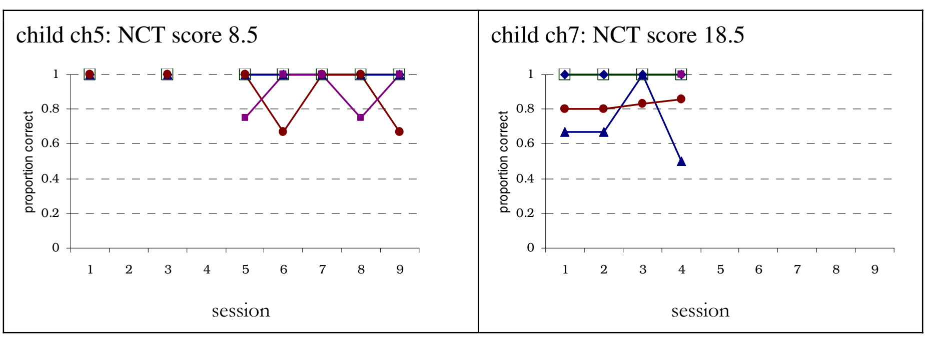
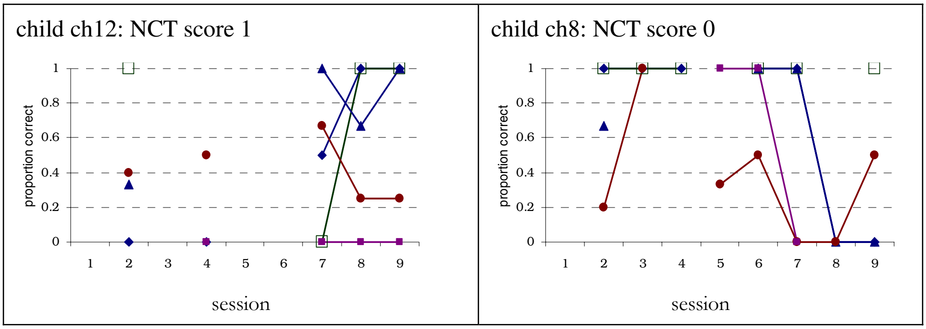


Table 4. For each session attended the proportion of t-stories answered correctly in each category by child 8 and child 12. For legend please see Table 2.



4 Discussion

Our results show that TouchStory can identify those aspects of narrative which individual children find difficult. Most of the children fall into three clusters; there is a correspondence between these clusters and the children's scores on the paper-based picture narrative comprehension task (NCT). The children in cluster 2 have the lowest NCT scores, in the range 0.5 to 2.5; those in cluster 1 are in the range 3 to 6.5, and those in cluster 3 are in the range 8.5 to 18.5. The children in cluster 1 show the clearest differentiation among proto-narrative categories, those in cluster 3 (who are very successful with TouchStory) show a similar but compressed differentiation. The children in cluster 2 show less differentiation. Interestingly cluster 2 includes both children who have diagnoses other than autism. Note, it is not our purpose to demonstrate that children with autism differ from those without; rather we take this result as an encouraging indication that our proto-narrative approach is particularly appropriate for children with autism.

Any improvement in the abilities of individual children with autism to deal with narrative in a way that illuminates their social world can only occur over an extended elapsed time. Therefore the results presented here will inform the design of a second longitudinal study which we plan for next year. The focus of the new study will be to investigate whether children with autism can *learn* about narrative from an adaptive interactive system such as TouchStory in a way which improves their understanding of stories in general. In particular we note that an adaptive system - which identifies those aspects of narrative which the child finds difficult and then adapts itself to focus on those aspects to provide an enjoyable but challenging game - is appropriate for children such as those in cluster 1. For children such as those in cluster 2, who present a more amorphous profile, we consider it might be more appropriate to focus on one aspect of narrative and investigate improvement in that one particular area. Although the children in cluster 3 perform similarly in the TouchStory study, their understanding of narrative differs greatly, their NCT scores differ, and their teacher's reports on their understanding of stories differ. This indicates that there are some aspects or complexities of narrative which TouchStory is not currently addressing. The next study will address these issues while preserving the emphasis on an enjoyable context.

Acknowledgements

We thank the children and staff involved in this study. Also, we thank René te Boekhorst, University of Hertfordshire, for guidance on statistical measures.

References

1. Frith, U., Explaining the Enigma. 1989: Blackwells.
2. Powell, S.D., Autism, in Developmental Psychology, D.J. Messer, Millar, S., Editor. 1999, Cambridge: Cambridge. p. 243-261.
3. Wing, L., The Autistic Spectrum: A Guide for Parents and Professionals. 1996, London: Constable.
4. Tager-Flushberg, H. and K. Sullivan, Attributing mental states to story characters: A comparison of narratives produced by autistic and mentally retarded individuals. Applied Psycholinguistics, 1995, 16: p. 241-256.
5. Capps, L., M. Losh, and C. Thurber, The Frog Ate the Bug and Made his Mouth Sad; Narrative Competence in Children with Autism. Journal of Abnormal Child Psychology, 2000. 18(2): p. 193-204.
6. Abell, F., F. Happe, and U. Frith, Do triangles play tricks? Attribution of mental states to animated shapes in normal and abnormal development. Cognitive Development, 2000 (15): p. 1-15.
7. Bruner, J. and C. Feldman, Theories of mind and the problem of autism, in Understanding other minds: perspectives from autism, S. Baron-Cohen, Editor. 1993, Oxford University Press: Oxford.
8. Dautenhahn, K., The Origins of Narrative: In Search for the Transactional Format of Narratives in Humans and Other Social Animals. International Journal of Cognition and Technology: Co-existence, Convergence, Co-evolution, 2002. 1(1): p. 97-123.
9. Hutto, D., Folk Psychological Narratives and the Case of Autism. Philosophical Papers, 2003. 32(3): p. 345-361.
10. Aurora. Aurora. 2000 [cited May 2004]; Available from: http://www.aurora-project.com.
11. Bruner, J., Actual Minds, Possible Worlds. 1986: Harvard University Press.
12. Paris, P. and P. Paris, Assessing Narrative Comprehension in Young Children. Reading Research Quarterly, 2003 38(1): p. 36-76.
13. Davis, M., et al., Towards an interactive system eliciting narrative comprehension in children with autism: A longitudinal study, in Designing Accessible Technology, P. Clarkson, P. Langdon, and P. Robinson, Eds. 2006, Springer-Verlag: London.

Computer-Assisted Language Learning (CALL) for Dyslexic Students

Cara Nicole Greene

School of Computing, National Centre for Language Technology, Dublin City University, Ireland
cgreene@computing.dcu.ie

Abstract. This paper outlines the analysis, design, development, deployment and evaluation stages of Computer-Assisted Language Learning (CALL) software aimed at dyslexic learners. CALL is traditionally aimed at second language acquisition. The research presented here is different because the target group are students with first language difficulties. Dyslexia is a specific learning disorder, which affects roughly eight percent of the population in Ireland [5]. This research identifies the lack of support for dyslexic teenagers in Irish secondary schools and establishes their particular needs. The paper describes CALL courseware development in progress, which aims to aid the reading, organisation and exam skills of dyslexic students through the use of online authentic Junior Certificate history curriculum texts. The final evaluation phase of the project will assess the efficacy of the software and investigate the question of whether dedicated software can improve the reading skills of teenage dyslexic students at word, sentence and text level.

1 Introduction

This paper identifies a lack of learning support for dyslexic teenagers in Irish secondary schools and provides a basis for catering to their needs through the use of curriculum-focused CALL comprehension software. Section 2 briefly introduces the state-of-the-art in research and application for dyslexic students. Section 3 discusses the research and methodological approach adopted. Section 4 presents the findings from the Analysis Phase of my research. Section 5 presents the Design phase output. Section 6 discusses future plans. This is followed by the conclusion and references.

2 State-of-the-Art in R&D and Application

Teaching experience and teacher and student surveys carried out within this research point to three main categories of tool-technology that are being used by dyslexic students in Irish schools: (I) general-purpose ICT tools (e.g. word processors), (II) generic speech processing tools (e.g. text-to-speech) and (III) special needs-focused tools (e.g. speech-recognition), which are usually expensive.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 793–800, 2006.

In some Irish secondary schools, probably as a result of a lack of funding, CALL materials for special needs primary school children are being used as a stopgap for young teenagers with learning difficulties because the linguistic content (lack of complexity) is deemed appropriate. However, the drawback to these materials is that the content is not age-appropriate for a teenager.

Tools that allow students to access text are most often used in secondary schools as opposed to dedicated CALL systems. Dragon Dictate [3] is a speech recognition system, which allows students to dictate a text, but it takes a long time to train the system to a student's voice and therefore it is often not very practical. Kurzweil [3] allows a student to scan in texts. However, both tools are very expensive and most schools do not have continuous access to them.

To date, there has not been much dedicated research and development in curriculum-focused design or development for dyslexic students. For example, Elkind et al. [4] report on the usefulness of text-to-speech systems to improve reading comprehension of students with dyslexia but their paper is more focused on the area of generic educational tools to help access text rather than providing a focused CALL environment. The Kar2ouche Shakespeare range [6] provides programs that present text and audio from Shakespeare's plays with the complete character set and a variety of backgrounds, props and sound effects. Some of these plays are options on the Junior Certificate English curriculum in Ireland.

3 Research and Methodological Approach

The research presented in this paper describes curriculum-focused comprehension software based on the Junior Cert history curriculum and popular texts, which uses authentic texts to give dyslexic teenagers more practise with the vocabulary they need in exams and socially. The research follows Colpaert's [1] Research-Based Research-Orientated (RBRO) design model as the theoretical basis underlying the overall research methodology. This model, which is a working hypothesis for pedagogy-driven, research-orientated development, takes as input all the facts, principles and requirements that the developer must take into account before starting to develop the application.

Colpaert's design methodology has its basis in the ADDIE (Analysis, Design, Development, Implementation and Evaluation) model grounded in the Waterfall model from Software Engineering [10]. This 'waterfall' model has successive and verifiable stages. Each step yields an output, which then serves as input for the next engineering phase. The idea is that each step is separate so that, for example, the design phase is reusable for other platforms because it is code-free.

Colpaert's Analysis phase entails gathering all possible information about relevant epistemological, empirical, actorial, contextual, technological, feasibility-related, and perceptive aspects, facts, findings, principles, and considerations that must be taken into account for the Design and subsequent phases.

Design is considered to be a working hypothesis, which can be verified and validated (or adjusted) after each implementation and evaluation. Each iteration of the design, implementation and evaluation loops leads to new working hypotheses. This process is the so-called engineering loop. The RBRO design model is divided into three stages: conceptualisation, specification and prototyping. Conceptualisation involves the identification of personas (learner-types from Analysis phase), the hypothesisation of practical goals, the formulation of scenarios, and the description of system tasks. Specification describes the back-end and the front-end. Prototyping is testing of the system and involves formative evaluation.

As the Design phase is quite comprehensive, this makes the Development phase easier. Development does not involve more design issues but is simply the execution of a plan conceived in the Design phase. In this way, educators play a more significant role in the CALL engineering process in the code-free Analysis and Design phases. Development involves a thorough testing process. The requirements set out in the Analysis and Design phases should guarantee an efficient deployment of the developed system during the Implementation phase into the targeted learning and teaching environment.

The Evaluation phase of Colpaert's model uses summative (after the Implementation stage) evaluation. Only after Implementation and Evaluation does the model call for a new working hypothesis for new versions or new applications, which will be developed based on extensive user feedback.

4 Analysis Phase Findings

This section discusses the research tasks carried out and the results established under the Analysis phase of Colpaert's model. The Analysis phase covers a literature review, needs analysis and matching CALL functionalities against dyslexic needs analysis.

4.1 Literature Review

The literature review consisted of a thorough analysis of each area relating to the research including existing courseware (see Section 2 for details), special education, government policy, curriculum analysis using Computational Linguistic (CL) technologies, Human-Computer Interaction (HCI) and usability issues, with the main focus on Computer-Assisted Language Learning (CALL) literature and Dyslexia.

CALL software is language-learning software (e.g., web-based, CD-Rom, interactive) that has lessons and exercises designed and developed for the particular needs of a target language learner group. CALL is a means of aiding the work done in the classroom by the teacher and can also be a means of independently learning a language. The literature review also involved a review of CALL design and development in both first and second language areas.

The word dyslexia is derived from the Greek 'dys' (meaning poor or inadequate) and 'lexis' (words or language). The word dyslexia therefore means 'difficulty with words'. The causes and characteristics of dyslexia were investigated, as well as the needs and difficulties of dyslexic students. Research in the area of dyslexia and software design for dyslexia was also reviewed.

4.2 Needs Analysis

The needs analysis of dyslexic teenagers was carried out through teaching experience, training with the Dyslexia Association of Ireland (DAI) and survey questionnaires of dyslexic teenagers and teachers. This research helped to identify the needs of dyslexic students and the learner types.

Needs analysis was carried out during teaching experience as a learning support and English as a Foreign Language (EFL) teacher (2002-03) in a secondary school. This teaching experience helped identify dyslexic student difficulties, needs and preferences. Professional dyslexia training with the DAI was an important component informing needs analysis e.g. ethical issues such as student privacy.

Learner types provide the 'personas' in Colpaert's [1] RBRO model. Learner-types were identified based on research into special education, dyslexia and discussions with the DAI. Feedback from the teacher, student and tutor surveys helped to identify these learner-types. Previous learning support teaching experience with students with learning difficulties and dyslexia proved valuable.

Survey questionnaires were carried out with special needs teachers (21 participants), dyslexia tutors (15 participants) and dyslexic teenagers (32 participants). The aim of the questionnaires was to investigate:

- What (if any) software/ICT tools students and teachers used for schoolwork?
- What type of software/tool teachers and students would like?
- Which features would be most beneficial and useful in this software/tool?

The results from the surveys show that the majority of special needs teachers and dyslexic students in schools use general-purpose ICT tools for their schoolwork as opposed to dedicated systems. Respondents said much of the software for dyslexic students is aimed at younger students and therefore not age-appropriate. Details of the ICT materials used in schools and current research in the area can be found in Section 2. The most often asked for (yet non-existent) software/tool from the surveys was comprehension software working with *authentic* texts. Respondents (teachers, tutors and students) pointed to dyslexic students' need for a new way to access the curriculum and to gain experience with exam and every-day language.

Respondents were presented with many features that can potentially be integrated into such a system. Table 1 in section 4.3, entitled "Key Features for Dyslexia-Focused CALL Software" shows the features chosen, how many respondents chose that feature and their reason from a dyslexic student need point of view. These results have been matched against how the CALL system could integrate this feature.

4.3 Matching CALL Functionalities Against Dyslexic Needs Analysis

This section takes the information gathered in my Analysis phase (following Colpaert's [1] model), which includes student needs, teacher needs, the survey questionnaire results and the review of CALL and related disciplines and puts them together in preparation for the Design phase of the model. Table 1 shows the key features for curriculum-focused comprehension software for dyslexic teenagers and how CALL software can cater to the needs of these students.

Table 1. Key Features for Dyslexia-Focused CALL Software

Feature	Who	Why	CALL?
Present text in digestible chunks	85% of learning support (LS) teachers	Short term memory and sequencing difficulties	Dyslexia-focused design and layout
Pictures, audio, video, TTS systems	96% of students & 100% of LS teachers	Poor symbol-sound meaning, grapheme phoneme linking	Multi-modal activities (audio & video)
Manipulate text like word processor	95% of LS teachers & 93% tutors	Defective fine motor skills, energy	Mind map, click & drag capabilities
Language quizzes & games	100% of students, 90% LS teachers, 100% tutors	Thinking in pictures, ADD, concentration	Colour, graphics, kinaesthetic features
Manual text summarisation	73% of dyslexia tutors	Organisation, STM, thinking in pictures	Multi-modal capabilities
Automatic text summarisation	26% of dyslexia tutors	Organisation, STM, thinking in pictures	Multi-modal capabilities
Authentic texts	93% of tutors	Need real texts	Exam materials
Memory games	46% of tutors	STM, sequencing	Multi-modal games

Due to Short-Term Memory (STM) and sequencing difficulties, students can lose their way in comprehension texts because they are not retaining important facts and therefore do not fully understand the text. CALL courseware's multi-modal capabilities can present a text in small digestible chunks. STM and defective phonological and visual access problems can make retaining grapheme-phoneme links difficult. CALL technology can strengthen the symbol-sound-meaning link with constant audio-visual revision because a student can click on a word and hear what it sounds like as often as they like.

CALL can address defective fine motor skills because it allows the student to manipulate text easily, summarise their work automatically and manually in mind-maps with pictures and record their voice instead of typing. This will also address organisation difficulties. CALL can address the problem of Attention Deficit Disorder (ADD). Many tasks that other students take for granted are difficult for dyslexic students due to the energy required to do it. Evidence from brain imaging suggests that while children with dyslexia do not activate the left hemisphere (language centre) in the brain as much when reading as non-dyslexic readers [9], they use five times of the overall brain area as non-dyslexic readers while performing a simple language task [8]. Students can lose interest and become frustrated reading a text. CALL programs can use colour, graphics, sound and kinaesthetic features to utilise multi-sensory paths to organisation and retrieval [12] rather than being restricted to the textual and linear format of textbooks. Games and quizzes can be integrated easily to give the student a new and interesting way of accessing the curriculum.

4.4 Curriculum Analysis Using CL Technologies

The content for my web-based CALL comprehension tool for dyslexic teenage students is provided by authentic texts from the History Junior Certificate curriculum. The content was determined in discussions with the DAI. Dyslexic students find the language in the history curriculum particularly difficult. These texts are online exam papers and online research samples of the History curriculum textbooks. This integrates the software with the curriculum in a blended CALL scenario and increases the reference of the software to students. Taggers, morphological analysis and pattern matching are used to analyse the text that students read to find words and patterns that dyslexic students find difficult, before the texts are presented to the students. Frequency lists are used to highlight words that occur repeatedly in these history texts e.g. 'archaeologist'. 'Hard words' and nouns that occur frequently in the history syllabus are stored in a XML database, which has associated pictures, audio files, and videos if appropriate. 'Hard words' are words of length >6, irregular phoneme-grapheme matching e.g. ph, words containing pre/suffixes, words that can be easily misinterpreted (p, b, d, q can all be seen as the same shape). There are also authentic popular texts such as texts from online teen magazines and online newspapers for the private newsgroup and pin board sections of the system.

The students will be able to avail of text-to-speech (TTS) systems (Talking Dictionary 8 [11] and Logox WebSpeech [7] and a summarisation system, Copernic Summariser [2]. These systems will be adapted to be used in my interactive environment so that students can access the text fully. Talking Dictionary 8 [11] is used because it is based on the Wordnet 2.1 [13] database and contains over 250,000 words together with audio information. Wordnet will provide synonyms/antonyms etc. in the multi-modal database and these words are also available in the Talking Dictionary.

5 Design Phase Findings

This section discusses the research tasks carried out under the Design phase following Colpaert's model and the associated findings and results.

5.1 Conceptualisation Results

Conceptualisation involved the identification of personas (learner-types from Analysis phase), the hypothesisation of practical goals, the formulation of scenarios, and the description of system tasks. Personas are types of users with common goals that have been identified in the Analysis phase. An example of one persona is 'John: John has an exemption from Irish. He avoids reading out loud in class and becomes annoyed and aggressive if pushed into reading'. Practical target goals for these personas were identified, such as improving reading skills, motivation, organisation and presentation skills. Scenarios were drawn up as to how personas will use the CALL tool. System tasks were described, which involved planning out exactly what each module in the system is going to do.

5.2 Specification Results

Specification entails the design of the back-end and the front-end of the system. The back-end of the system describes the system structure in terms of components and their interaction. The front-end describes what the students will see and how they interact with the system.

Front-End of the System. The student logs into the system and they can go into history work, schoolwork/student newsgroups or schoolwork/student pin boards. When the student enters the history section the appropriate level of texts, decided by the teacher/tutor, are displayed. Each text will be presented in small digestible chunks (including text, recorded voice, pictures, video) to ease the load on the student's STM. Students will be able to click on words to hear them and access extra material: audio, text, video, antonyms, and synonyms. The student will be able to work on the texts and associated exercises at their own pace. The constant audio-visual revision will strengthen the symbol-sound-meaning (grapheme-phoneme) link for the students. The student will be able to manually summarise the text by clicking-and-dragging words from the text, and pictures and related content from the database of history nouns into a mind-map area for themselves, as well as record oral summaries. The teacher or tutor will have a login so that they can see students progress throughout the system.

Back-End of the System. The back-end consists of the XML database that contains all the information for the comprehension texts and related information. The system has two main areas: schoolwork and a student area. The system modules are all separate to allow content to be added easily. The small texts that students receive have already been pre-processed in the Analysis phase and are stored in an XML database with associated video, audio and text. Further information can be found in Section 4.4.

5.3 Prototyping Results

The purpose of prototyping is to test discrete functionalities. Prototyping involves testing sections and versions of the software on a number of different people: fellow developers, teachers and finally when the product is ready, the students, for the implementation and evaluation cycle. Functionalities such as the XML database, summarisation and text-to-speech have been tested.

6 Future Plans - Development, Deployment, Evaluation

Formative evaluation of discrete functions within the system is under way. The prototype is being developed at the moment. The software will then be deployed with students who receive after-school dyslexia tutoring to help them with their school and exam work. This first cycle of implementation and summative evaluation will provide student and teacher feedback, which will feed back into the design phase and create new hypotheses for further development. These changes will then be implemented and further summative evaluation carried out.

7 Conclusion

The paper provides a background to research and development in the area of learning software for teenage dyslexic students. My research identifies and focuses on the lack of support for dyslexic teenagers in Irish secondary schools and establishes the particular needs of this age group. The paper describes CALL courseware development in progress. The software aims to aid the reading, organisation and exam skills of dyslexic students through the use of online authentic Junior Certificate History curriculum texts. The final evaluation phase of the project will assess the efficacy of the software and help answer the question of whether dedicated software can improve the reading skills of teenage dyslexic students at word, sentence and text level.

References

1. Colpaert, J.: Design of Online Interactive Language Courseware: Conceptualization, Specification and Prototyping. Research into the impact of linguistic-didactic functionality on software architecture. PhD thesis, University of Antwerp, Belgium (2004)
2. Copernic Summariser: Copernic Summariser. Available at `http://www.copernic.com` (2005)
3. Edtech: Edtech Educational Software. Available at `http://www.edtech.ie` (2005)
4. Elkind, J., Cohen, K., Murray, C.: Using Computer-Based Readers to Improve Reading Comprehension of Students with Dyslexia. Annals of Dyslexia, v43. (1993) 238–259
5. Holden, L.: Dealing with Dyslexia. Provided by the Irish Times. Available at `http://skoool.ie/skoool/parents.asp?id=2912` (2004)
6. Kar2ouche: Kar2ouche Romeo and Juliet. Available at `http://www.immersiveeducation.com` (2006)
7. Logox Webspeech: Logox Webspeech. Available at `http://www.webspeech.de/index2.php` (2003)
8. Richards, T., Dager, S., Corina, D., Serafini, S., Heide, A., Steury, K., Strauss, W., Hayes, C., Abbott, R., Craft, S., Shaw, D., Posse, S., Berninger, V.: Dyslexic Children Have Abnormal Brain Lactate Response to Reading-Related Language Tasks. American Journal of Neuroradiology, 20. (1999) 1393–1398
9. Serafini S, Steury K, Richards T, D, Abbott R, SR, Berninger V.: Comparison of fMRI and PEPSI during language processing in children. Magn Reson Med, 45. (2001) 217–225
10. Sommerville, I.: Software Engineering. Harlow, Essex, England, New York: Pearson/ Addison Wesley. 7th Edition, 2004
11. Talking Dictionary 8: Talking Dictionary 8. Available at `http://www.brothersoft.com` (2003)
12. Wahl, R.: The Role of Computers in Teaching EFL to Dyslexic Students. ETNI. Available at `http://www.etni.org.il/etninews/paper.htm` (1996)
13. Wordnet 2.1: Available at `http://wordnet.princeton.edu/obtain` (2005)

A New Audio Testing System for the Newly Blind and the Learning Disabled to Take the National Center Test for University Admissions

Mamoru Fujiyoshi[1] and Akio Fujiyoshi[2]

[1] Research Division, the National Center for University Entrance Examinations
2-19-23 Komaba, Meguro, Tokyo, 153-8501, Japan
fujiyosi@rd.dnc.ac.jp

[2] Department of Computer and Information Sciences, Ibaraki University
4-12-1 Nakanarusawa, Hitachi, Ibaraki, 316-8511, Japan
fujiyosi@mx.ibaraki.ac.jp

Abstract. A new audio testing system was developed for the newly blind and the learning disabled (dyslexia) who have difficulties with reading braille or print-format tests. The system enables them to take the National Center Test for University Admissions. The system was developed primarily on a tablet PC and presents not only speech sound but also document structures and figures from the test.

1 Introduction

The National Center Test for University Admissions is the joint achievement test for admissions into all national and local public universities as well as many private universities in Japan. Every year, about 600,000 students take the National Center Test. As for test-takers with disabilities, special arrangements regarding test media such as large print test and braille-format test have been administered [7]. Audio tests, however, have not been administered yet. There are difficulties in administering ordinary types of audio tests for the National Center Test because the documents are very long and the document structure very complicated. This study contributes to the development of a new audio testing system for the newly blind and the learning disabled (dyslexia) who have difficulties with reading braille or print-format tests.

Ordinarily, audio tests are administrated by means of readers [4], audio cassettes [2,9,10], DAISY (Digital Audio Accessible Information System) [3] or computers [1]. The simplest procedure is to recruit readers and have them read out a test booklet to a test-taker directly, but it is not easy to find enough well-trained readers for each test-taker. And for fairness and security reason, it might be necessary to supervise such readers by another person. Audio cassettes make it easy for test-takers to listen to the test sequentially, but it is inconvenient to go directly to a particular section of the test unless rewinding and fast-forwarding can be done easily. Recently, it has become possible to design audio tests by means of DAISY, which is a world standard audio system for people with

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 801–808, 2006.

visual disabilities, taking the place of audio cassettes. DAISY offers audio tests in CD quality, and test-takers can listen to the test from any point, such as from an underlined or blank part, without delay. They can also use the talk-speed-control function, by which the speech sound can be adjusted from 1/2 to 3 times normal speed. As a result of the experiments with the Law School Aptitude Test [6] and the National Bar Examination [5] of Japan, we found that the DAISY tests were almost equal to braille-format tests for the blind in design and administration. However, DAISY is not convenient enough for tests which have complicated document structure. A computer test is also inappropriate for tests written in Japanese even with advanced screen-reader software for the blind because of the ambiguity of the reading of Kanji in Japanese sentences. Screen-reader software can not always convert Japanese sentences into correct Japanese speech.

We developed a new audio testing system utilizing a tablet PC which enables the newly blind and the learning disabled to take the National Center Test for University Admissions. This system presents the document structure of test problems with characters or icons on a computer screen for the learning disabled and on braille paper for the newly blind. Test-takers can take a variety of tests by using the electronic pen of a tablet PC.

This paper is organized as follows: Section 2 contains a summary of the system; Section 3 provides a system evaluation by experiment; and Section 4 is the conclusion.

2 The Tablet PC Audio Testing System

2.1 Composition

We developed a new audio testing system utilizing a tablet PC. The composition of the system for the newly blind is slightly different from the one for the learning disabled.

For the newly blind, the composition of the system is illustrated in Fig. 1. The system consists of a PC (FMV STYLISTIC, Fujitsu Inc.), an A3 tablet (Intuos3, Wacom Inc.), a ten-key pad (NT-1U, Sanwa Supply Inc.), and a speaker with amplifier (SRS-T88, Sony Corp.). On the A3 tablet, a sheet of braille paper, on which the document structure for each problem or figure has been embossed by a braille printer, is mounted. With just an electronic pen, test-takers can listen to speech sound from any part of the document and answer the questions. The ten-key pad is used to adjust the speed and volume of speech sound.

For the learning disabled, the system consists only of a tablet PC (VAIO PCV-LX80, Sony Corp.). On the LCD (Liquid Crystal Display) screen of the tablet PC, the document structure is displayed. Speech sound comes from speakers mounted on both sides of the LCD screen.

The software was exclusively developed at the National Center for University Entrance Examinations to run on Windows XP Tablet PC Edition with Visual C++ 6.0 and Windows Media Player SDK 10 (Microsoft Corp.).

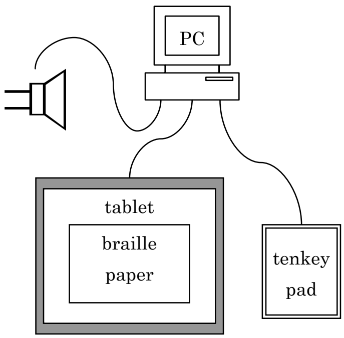


Fig. 1. The composition of the system for the newly blind

2.2 Usage

The system displays the document structure of each problem, which enables a test-taker to confirm the document structure throughout the answering process and to point to any position of document to listen to the speech sound.

Fig. 2 is an example of the document structure displayed on the LCD screen for the learning disabled. For the newly blind, on the other hand, only the character part of the document structure was embossed by a braille printer, which is mounted on the tablet. In Fig. 2, the first line shows the subject name 'Gensha 1' and the problem number 'Q1'. The upper part shows the document structure of the theme document of the problem. Each line 'p1'-'p4' corresponds to a paragraph in the theme document. The symbol 's' represents a sentence in a paragraph, and the symbols 'a'-'g' represent underlined parts of the theme document. The lower part shows the document structure of the questions of the problem. Each line 'q1'-'q7' corresponds to a question. The symbols '[]1'-'[]7'

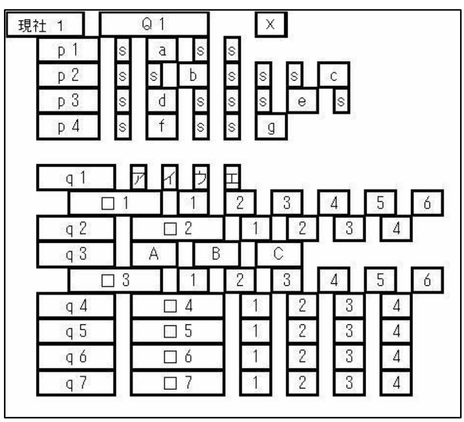


Fig. 2. An example of the document structure displayed on the LCD screen

represent answer items, and the symbols '1'-'6' represent possible answers for an answer item.

To produce speech sound, a test-taker just touches the symbol of either a paragraph, a sentence, an underlined part of theme document, an answer item or a possible answer with the electronic pen. To stop the sound, a test-taker touches the blank part of the screen with the electronic pen.

To answer a question, the tail switch of the electronic pen is used. When a test-taker successively touches the symbol of an answer item and the symbol of a possible answer with the tail switch of the electronic pen, an answer is recorded by the system, and a confirmation message is produced, for example, "On answer item No. 1, choice No. 2 was selected."

The ten-key pad is used to start the test, to suspend the test, to end the test, and to adjust the speed and volume of speech sound. When the ten-key pad is used, a confirmation message is produced.

The answering process of a test-taker is automatically recorded with the time whenever the electronic pen or the ten-key pad is used.

3 Evaluation Experiment

In order to evaluate the tablet PC audio testing system, an experiment was conducted by comparing the tablet PC audio test with other test media such as the braille-format test, the DAISY test, and the multimodal test of braille-format and tablet PC audio test.

3.1 Method

Two subject groups were assigned in this experiment, namely, a blind group and a non-disabled group. The blind group consists of 12 students from a high school for the blind, and the non-disabled group, 28 average high school students.

The experimental design for the blind group was a repeated 4x4 Latin-square method [8] because we could not use the same problem in different test media for the same person. The image of the experimental design for the Latin-square method is shown on Table 1. There were 4 test media: the braille-format test, the DAISY test, the tablet PC audio test, and the multimodal test for braille-format and tablet PC audio test. There were 4 subject groups, i.e., the blind group was evenly divided into 4 subgroups. Four problems were prepared from tests in "Contemporary Social Studies" previously used in the National Center Test. The allotment of marks, number of characters, size of figures, and quantity of problems are shown on Table 2. The size of a figure is converted to a number of characters (the density of characters in the document $\times$ the area of the figure).

The test procedure was administered without time limits. Blind subjects took the braille-format test by braille booklets, the DAISY test by a DAISY player (Plextalk Portable Recorder PTR1, Plextor Inc.), the tablet PC audio test by our system, and the multimodal test by using both the braille booklets and the tablet PC audio testing system. The behavior of blind subjects was observed by

Table 1. Image of the experimental design for the Latin-square method

		Subect Groups			
		Group 1	Group 2	Group 3	Group 4
Test Media	Braille	Problem 1	Problem 2	Problem 3	Problem 4
	DAISY	Problem 2	Problem 1	Problem 4	Problem 3
	Tablet	Problem 3	Problem 4	Problem 1	Problem 2
	Multi	Problem 4	Problem 3	Problem 2	Problem 1

Table 2. Allotment of marks, number of characters, size of figures, and quantity of problems

	Marks	Characters	Figures	Quantity
Problem 1	20	3550		3550
Problem 2	20	3314	1196	4510
Problem 3	20	4374		4374
Problem 4	20	4141	177	4318
Total	80	15379	1373	16752

test monitors, and the answer-process time of blind subjects was recorded by the monitors using stop watches. On the other hand, the non-disabled subjects answered the same 4 problems using print-format tests. They recorded their own time using stop watches.

3.2 Result

Score Distribution: For the blind group, the score distribution of the tablet PC audio test was almost the same as those of the braille-format test, the DAISY test and the multimodal test. The Box-and-whiskers plots [8] in Fig. 3 represent the score distribution of each test media for the blind group, the distribution of the total score of the four test media for the blind group, and the distribution of the total score of the print-format test for the non-disabled group. The vertical lines in the middle of the box plots indicate the median. The '+' symbols in the boxes are the mean. Table 3 is the result of the analysis of the variance of the score using the repeated Latin-square design [8]. There were no significant main effects on score regarding all three factors: test media, subject group and problems. The three-dimensional interaction was not significant.

Though the median of the total score for the non-disabled group was about 6 marks out of 80 higher than that for the blind group, there was no significant difference compared to the Mann-Whitney test [8].

Distribution of the Answering Speed: For the blind group, the distribution of the answering speed for tablet PC audio test was also similar to those for the braille-format test, the DAISY test and the multimodal test. The Box-and-whiskers plots in Fig. 4 represent the distribution of the answering speed of each test media for the blind group, the distribution of the total answering speed of the

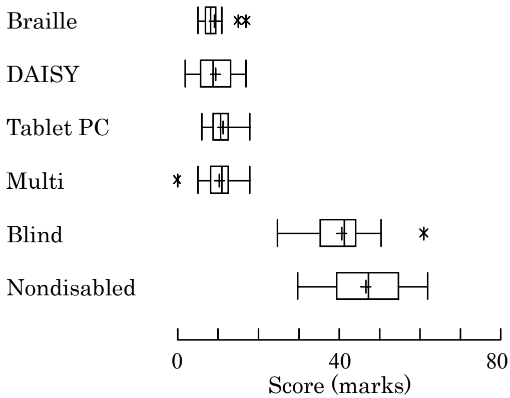


Fig. 3. Box-and-whisker plots for the score

Table 3. Result of the analysis of the variance of the score using the Latin-square design

Source	DF	Sum of Square	Mean Square	F Value	Pr > F
Group	3	59.165	19.722	1.05	0.383
Media	3	41.379	13.793	0.74	0.538
Problem	3	112.716	37.572	2.01	0.133
Group×Media×Problem	6	39.067	6.511	0.35	0.906
Error	32	599.415	18.732		
Corrected Total	47	851.741			

four test media for the blind group, and the distribution of the total answering speed of print-format test for the non-disabled group. Here, answering speed is defined as the number of characters of problems ('quantity' in Table 2) per minute to answer the problems for each person. Table 4 is the result of the analysis of the variance of the answering speed for the repeated Latin-square design. There were no significant main effects on answering speed regarding all three factors: test media, subject group and problems. The three-dimensional interaction among the three factors was not significant.

We found that the median of the total answering speed for the non-disabled group was about 2.4 times faster than that for blind subject group. There was a significant difference in the distribution of the total answering speed between the blind group and the non-disabled group according to the Mann-Whitney test ($p < 0.0001$).

3.3 Discussion

The result of the evaluation experiment shows that the tablet PC audio test and the DAISY test are almost equal to the braille-format test for the blind in terms of score and answering speed without time limits. We also found that the

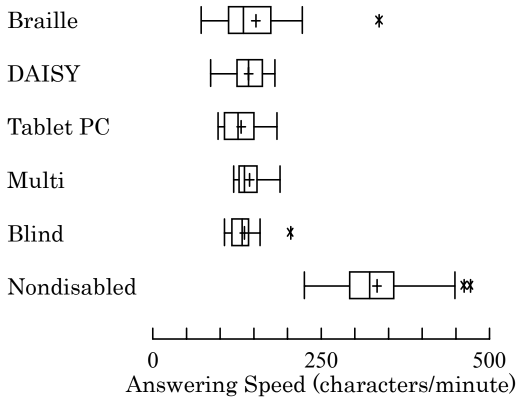


Fig. 4. Box-and-whisker plots for the answering speed

Table 4. Result of the analysis of the variance on the answering speed using the Latin-square design

Source	DF	Sum of Square	Mean Square	F Value	Pr > F
Group	3	3457.613	1152.538	0.57	0.640
Media	3	3013.639	1004.546	0.50	0.688
Problem	3	4256.247	1418.749	0.70	0.559
Group×Media×Problem	6	6435.548	1072.591	0.53	0.782
Error	32	64877.108	2027.410		
Corrected Total	47	82040.156			

answering speed of the print-format test for non-disabled test-takers is about 2.4 times faster than of the four test media for blind test-takers provided that both have the same achievement level. These findings are consistent with our previous results for the DAISY test in the Law School Aptitude Test [6] and the National Bar Examinations [5].

The tablet PC audio testing system can be introduced to the National Center Test if the system is improved. Though the blind subjects were not trained well, they could take the tablet PC audio test with similar scoring ability and answering speed as with the braille-format test.

With time limits, audio test may be more useful because audio test such as the tablet PC audio test and the DAISY test have a talk-speed-control function. The tablet PC audio test may be a more effective test media than the DAISY test when the document structure is complicated.

The tablet PC audio testing system enables the newly blind and the learning disabled with reading difficulties to take tests effectively. It is, however, advisable that blind test-takers without reading difficulties choose test media from among audio tests and a braille-format test because they have studied with braille textbooks and taken braille-format tests in quizzes and term examinations.

4 Conclusion

We developed this new audio testing system utilizing a tablet PC for the newly blind and the learning disabled who have difficulty reading long documents. In most advanced countries, audio tests are prepared for the newly blind and the learning disabled [2,9,10]. Ordinarily, they are administered with readers, audio cassettes, DAISY players or computers. Those methods, however, are inappropriate for the National Center Test for University Admissions because of complicated document structure. Therefore, we designed this new audio testing system that presents the document structure of a test.

As a result of the experiment for evaluation, we found that a tablet PC audio test was almost equal to a braille-format test for the blind in terms of score and answering speed.

To put the tablet PC audio testing system to practical use, we still have two challenges: improvement of the management software for the electronic pen; and development of an authoring system for the tablet PC audio tests so that anyone can easily produce material for audio tests.

References

1. Allan, M. J., Bulla, N. and Goodman, S. A.: Test access: guidelines for computer administered testing. American Printing House for the Blind, Kentucky (2003)
2. Allman, C. B.: Test access: making tests accessible for students with visual impairments: A guide for test publishers, test developers, and state assessment personnel, 2nd ed. American Printing House for the Blind, Kentucky (2004)
3. DAISY Consortium: DAISY 2.02 Specification. `http://www.daisy.org/ publications/specifications/daisy_202.html` (2001)
4. Educational Testing Service: Resources for test takers with disabilities: Guidelines for a Test Reader. `http://www.ets.org/disability/index.html` (2006)
5. Fujiyoshi, M. and Fujiyoshi, A.: Universal design for multiple-choice tests in secondary stage of National Bar Examination: estimation of extension ratio of testing time and design for audio format test. Research Journal of University Entrance Examinations (in Japanese) 15 (2005) 27-34
6. Fujiyoshi, M.: Universal design for Law School Aptitude Test: design for digital audio format test and braille format test. Research Journal of University Entrance Examinations (in Japanese) 14 (2004) 15-24
7. Fujiyoshi, M. and Fujiyoshi, A.: Estimating testing time extension ratios for students with disabilities from item cumulative curves. New Developments in Psychometrics, Proceedings of the International Meeting of the Psychometric Society IMPS2001 (2003) 265-272
8. Porkees, R.: Collins dictionary of statistics. Harper Collins Publishers, New York (2004)
9. Ragosta, M. and Wendler, C.: Eligibility issues and comparable time limits for disabled and nondisabled SAT examinees. ETS Research Report, RR-92-35 (1992) 1-33
10. Willingham, W. W., Ragosta, M., Bennett, R. E., Braun, H., Rock, D. A. and Powers, D. E.: Testing handicapped people. Allyn and Bacon, Boston (1988)

Embodied Agents in Language Learning for Children with Language Challenges

Dominic W. Massaro

Department of Psychology
University of California, Santa Cruz
Santa Cruz, CA 95064 U.S.A.
Massaro@ucsc.edu
http://mambo.ucsc.edu/dwm

Abstract. Given the value of face-to-face interaction in communication and learning, our persistent goal has been to develop, evaluate, and apply animated agents to produce realistic and accurate speech. We have implemented these agents as computer-assisted speech and language tutors for hard of hearing and autistic children, and other children with language challenges. Our language-training program utilizes conversational agents, who guide students through a variety of exercises designed to teach vocabulary and grammar, to improve speech articulation, and to develop linguistic and phonological awareness. We report a new experiment showing its effectiveness for school children learning English as a new language. Some of the advantages of this pedagogy and technology include the popularity and effectiveness of computers and embodied conversational agents, the perpetual availability of the program, and individualized instruction. Animated tutors offer a promising approach to language learning, human-machine interaction, and education.

1 Introduction

Language challenges are pervasive in today's world. Given the maximum migration that our society has ever experienced, many individuals become members of a new linguistic community. Given the highly mobile society, individuals of all walks of life find themselves in situations in which successful education, business, and social interactions require use of a nonnative language. As an obvious example, English is becoming increasing necessary and desirable, and the number of people in the world who are learning English is increasing at a rapid rate. Learning a new language, however, is a significant challenge for all individuals, whether young or old. In addition, there are surprisingly many individuals who have language and speech disabilities, and these individuals like those learning a new language require additional instruction in language learning. Currently, however, these needs are not being met because there are not enough skilled teachers and professionals to give them the one on one attention that they need. So they resort to other resources, such as books or other media, but the problems with these are that they are not easily personalized to the students' needs, they lack the engaging capability of a teacher, they are rather expensive, and they are relatively ineffective.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 809 – 816, 2006.

In this paper, we describe several language learning applications with a virtual tutor for language learning and speech training. We begin by describing research that demonstrates that our perception and understanding of language are influenced by a speaker's face and accompanying gestures, as well as the actual sound of the speech.

1.1 State of the Art: Value of the Face in Communication

Although traditionally speech has been viewed as solely an auditory phenomenon, speech as a multimodal phenomenon is supported by experiments indicating that our perception and understanding are enhanced by a speaker's face and accompanying gestures, as well as the actual sound of the speech [1,2]. There are several reasons why the use of auditory and visual information together is so successful. These include a) robustness of visual speech, b) complementarity of auditory and visual speech, and c) optimal integration of these two sources of information. Speechreading, or the ability to obtain speech information from the face, is robust in that perceivers are fairly good at speechreading even when they are not looking directly at the talker's lips [1]. Furthermore, accuracy is not dramatically reduced when the facial image is blurred (because of poor vision, for example), when the face is viewed from above, below, or in profile, or when there is a large distance between the talker and the viewer.

Complementarity of auditory and visual information simply means that one of the sources is strong when the other is weak. A distinction between two segments robustly conveyed in one modality tends to be relatively ambiguous in the other modality. Two complementary sources of information make their combined use much more informative than would be the case if the two sources were non-complementary [1].

Perceivers combine or integrate the auditory and visual sources of information in an optimally efficient manner. Many different empirical results have been accurately predicted by a Fuzzy Logical Model of Perception (FLMP) that describes an optimally efficient process of combination [1].

1.2 Embodied Conversational Agents

The value of visible speech in face-to-face communication was the primary motivation for the development of Baldi®, a 3-D computer-animated talking head. Baldi provides realistic visible speech that is almost as accurate as a natural speaker [1,3]. The goal of the visible speech synthesis carried out in the Perceptual Science Laboratory (PSL) has been to develop a polygon (wireframe) model with realistic motions (but not to duplicate the musculature of the face). Our animated face can be aligned with either the output of a speech synthesizer or natural auditory speech [3]. We have also developed the phoneme set and the corresponding target and coarticualtion values to allow synthesis of several other languages. These include Spanish (Baldero), Italian (Baldini), Mandarin (Bao), Arabic (Badr), French (Balduin), German (Balthasar), and Russian (Balda). Baldi, and his various multilingual incarnations are seen at: `http://mambo.ucsc.edu/psl/international.html`.

In our facial animation algorithm, each segment is specified with a target value for each facial control parameter. Coarticulation, defined as changes in the articulation of a speech segment due to the influence of neighboring segments, is based on a model of speech production using rules that describe the relative dominance of the characteristics of the speech segments [1,3].

A central and somewhat unique quality of our work is the empirical evaluation of the visible speech synthesis, which is carried out hand-in-hand with its development. The quality and intelligibility of Baldi's visible speech has been repeatedly modified and evaluated to accurately simulate naturally talking humans [1,3]. The gold standard we use is how well Baldi compares to a real person. Given that viewing a natural face improves speech perception, we determine the extent to which Baldi provides a similar improvement. We repeatedly modify the control values of Baldi in order to meet this criterion. We modify some of the control values by hand and also use data from measurements of real people talking.

Several advantages of utilizing a computer-animated agent as a language tutor are clear, including the popularity of computers and embodied conversational agents. Computer-based instruction is emerging as a prevalent method to train and develop vocabulary knowledge for both native and second-language learners and individuals with special needs [4]. An incentive to employing computer-controlled applications for training is the ease in which automated practice, feedback, and branching can be programmed. Another valuable component is the potential to present multiple sources of information, such as text, sound, and images in parallel. Instruction is always available to the child, 24 hours a day, 365 days a year. Furthermore, instruction occurs in a one-on-one learning environment for the students. Applications with animated tutors perceived as supportive and likeable will engage foreign language and ESL learners, reading impaired, autistic and other children with special needs in face-to-face computerized lessons. We now review several different applications utilizing Baldi to carry out language tutoring, and then report a new study in the learning of vocabulary in a new language.

2 Pedagogy of Language Learning

Vocabulary knowledge is critical for understanding the world and for language competence in both spoken language and in reading. There is empirical evidence that very young children more easily form conceptual categories when category labels are available than when they are not [4]. Even children experiencing language delays because of specific language impairment benefit once this level of word knowledge is obtained. It is also well-known that vocabulary knowledge is positively correlated with both listening and reading comprehension [5]. It follows that increasing the pervasiveness and effectiveness of vocabulary learning offers a timely opportunity for improving conceptual knowledge and language competence for all individuals, whether or not they are disadvantaged because of sensory limitations, learning disabilities, or social condition.

Learning and retention are positively correlated with the time spent learning. Our technology offers a platform for unlimited instruction, which can be initiated whenever and wherever the child and/or mentor chooses. Instruction can be tailored exactly to the student's need, which is best implemented in a one-on-one learning environment for the students. Other benefits of our program include the ability to seamlessly meld spoken and written language, and provide a semblance of a game-playing experience while actually learning. Given that education research has shown that children can be taught new word meanings by using direct instruction methods [5], we implemented these basic features in an application to teach vocabulary and grammar.

2.1 Research and Methodological Approach

To test the effectiveness of vocabulary instruction using an embodied conversational agent as the instructor, we developed a series of lessons that encompass and instantiate the developments in the pedagogy of how language is learned, remembered, and used.

One of the principles of learning that we exploit most is the value of multiple sources of information in perception, recognition, learning, and retention. An interactive multimedia environment is ideally suited for learning [4]. Incorporating text and visual images of the vocabulary to be learned along with the actual definitions and sound of the vocabulary facilitates learning and improves memory for the target vocabulary and grammar. Many aspects of our lessons enhance and reinforce learning. For example, the existing program makes it possible for the students to 1) Observe the words being spoken by a realistic talking interlocutor, 2) Experience the word as spoken as well as written, 3) See visual images of referents of the words, 4) Click on or point to the referent or its spelling, 5) Hear themselves say the word, followed by a correct pronunciation, and 6) Spell the word by typing, and 7) Observe and respond to the word used in context.

2.2 Effectiveness for Hearing Loss

It is well known that children with hearing loss have significant deficits in both spoken and written vocabulary knowledge. To assess the learning of new vocabulary, we carried out an experiment based on a within-student multiple baseline design where certain words were continuously being tested while other words were being tested and trained [6]. Although the student's instructors and speech therapists agreed not to teach or use these words during our investigation, it is still possible that the words could be learned outside of the learning context. The single student multiple baseline design monitors this possibility by providing a continuous measure of the knowledge of words that are not being trained. Thus, any significant differences in performance on the trained words and untrained words can be attributed to the training program itself rather than some other factor.

We studied eight children with hearing loss, who needed help with their vocabulary building skills as suggested by their regular day teachers. The experimenter developed a set of lessons with a collection of vocabulary items that was individually composed for each student. Each collection of items was comprised of 24 items, broken down into 3 categories of 8 items each. Three lessons with 8 items each were made for each child. Images of the vocabulary items were presented on the screen next to Baldi as he spoke. Assessment was carried out on all of the items at the beginning of each lesson. It included identifying and producing the vocabulary item without feedback. Training on the appropriate word set followed this testing.

As expected, identification accuracy was always higher than production accuracy. This result is expected because a student would have to know the name of an item to pronounce it correctly. There was little knowledge of the test items without training, even though these items were repeatedly tested for many days. Once training began on a set of items, performance improved fairly quickly until asymptotic knowledge was obtained. This knowledge did not degrade after training on these words ended and training on other words took place. In addition, a reassessment test given about 4 weeks after completion of the experiment revealed that the students retained the items that were learned.

2.3 Effectiveness for Autism

The tutoring application has also been used in evaluating vocabulary acquisition, retention and generalization in children with autism [7]. Although the etiology of autism is not known, individuals diagnosed with autism must exhibit a) delayed or deviant language and communication, b) impaired social and reciprocal social interactions, and 3) restricted interests and repetitive behaviors. The language and communicative deficits are particularly salient, with large individual variations in the degree to which autistic children develop the fundamental lexical, semantic, syntactic, phonological, and pragmatic components of language. Vocabulary lessons were constructed, consisting of over 84 unique lessons with vocabulary items selected from the curriculum of two schools. The participants were eight children diagnosed with autism, ranging in age from 7-11 years.

The results indicated that the children learned many new words, grammatical constructions and concepts, proving that the application provided a valuable learning environment for these children. In addition, a delayed test given more than 30 days after the learning sessions took place showed that the children retained over 85% of the words that they learned. This learning and retention of new vocabulary, grammar, and language use is a significant accomplishment for autistic children.

Although all of the children demonstrated learning from initial assessment to final reassessment, it is possible that the children were learning the words outside of our learning program (for example, from speech therapists or in their school curriculum). Furthermore, it is important to know whether the vocabulary knowledge would generalize to new pictorial instances of the words. To address these questions, a second investigation used the single subject multiple probe design, as was done in [6]. Once a student achieved 100% correct, generalization tests and training were carried out with novel images. The placement of the images relative to one another was also random in each lesson. Assessment and training continued until the student was able to accurately identify at least 5 out of 6 vocabulary items across four unique sets of images.

Although performance varied dramatically across the children and across the word sets during the pre-training sessions, training was effective for all words sets for all children. Given training, all of the students attained our criterion for identification accuracy for each word set and were also able to generalize accurate identification to four instances of untrained images. The students identified significantly more words following implementation of training compared to pre-training performance, showing that the program was responsible for learning. Learning also generalized to new images in random locations, and to new interactions outside of the lesson envirronment. These results show that our learning program is effective for children with autism, as it is for children with hearing loss.

We were gratified to learn that the same application could be used successfully with both autistic children and children with hearing loss [4]. Specific interactions can be easily modified to accommodate group and individual differences. For example, autistic children are much more disrupted by negative feedback, and the lesson can be easily designed to instantiate errorless learning. Given these successes, we carried out a study with English language learners (ELL).

2.4 Effectiveness for English Language Learners (ELL)

This study was carried out with English Language Learners (ELL) and involved the use of a recently-released application, Timo Vocabulary (`http://animatedspeech.com`), which instantiated the pedagogy we found in our earlier research 6,4,7]. Nine children ranging in age from 6-7 years were tested in the summer before first grade. Almost all of the children spoke Spanish in the home. The children were pretested on lessons in the application in order to find three lessons with vocabulary that was unknown to the children. A session on a given day included a series of three test lessons, and on training days, a training lesson on one of the three sets of words. Different lessons were necessarily chosen for the different children because of their differences in vocabulary knowledge. As shown in Figure 1, the test session involved the presentation of the images of a given lesson on the screen with Timo's request to click on one of the items, e.g., Please click on the oven. No feedback was given to the child. Each item was tested once in two separate blocks to give 2 observations on each item. Three different lessons were tested, corresponding to the three sets of items used in the multiple baseline design. A training session on a given day consisted of just a single lesson in which the child was now given feedback on their response. Thus, if Timo requested the child to click on the dishwasher and the child clicked on the spice rack, Timo would say, "I asked for the dishwasher, you clicked on the spice rack. This is the dishwasher. The training session also included the Elicitation and Imitation sections in which the child was asked to repeat the word when it was highlighted and Timo said it, and the child was asked to say the item that was highlighted. Several days of pretesting were required to find lessons with unknown vocabulary. Once the 3 lessons were determined, the pretesting period was followed by the training days. Given that the children learned at different rates, the results were averaged with respect to when the different training regimens were initiated. Thus, for example, the results at block 3 give performance before training on Set 1 words, and the results at block 4 give performance after one block of training on Set 1 words. As can be seen in Figure 1, training was effective in teaching new vocabulary. Some of the variability in the figure is due to having just a few subjects at some blocks: Blocks 9 and 15 have just 1 subject's data.

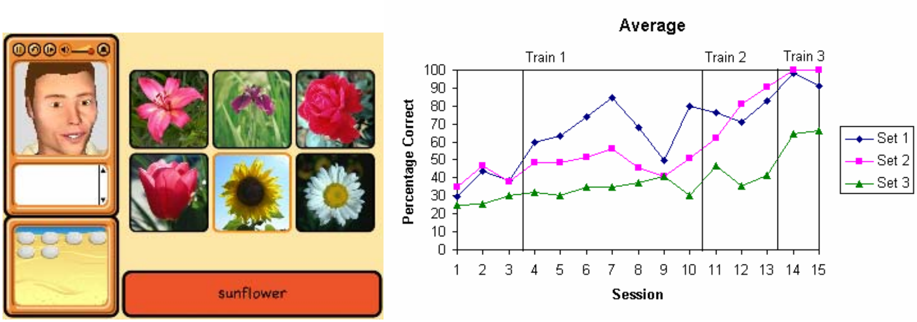


Fig. 1. The left panel shows a typical screen shot from the Timo Vocabulary application. The right panel gives the average percentage of correct identifications for each of the three sets of words. The three vertical bars indicate when training was initiated for each of the three sets, respectively. Performance improved on each set of words after training on that set was initiated.

3 Future Plans and Conclusion

Animated Speech is releasing a Lesson Creator that allows easy creation of new language lessons. This user-friendly application allows the composition of lessons with minimal computer experience and instruction. Although it has many options, the program has wizard-like features that direct the coach to explore and choose among the alternative implementations in the creation of a lesson. The application will include a curriculum of thousands of vocabulary words, and can be implemented to teach both individual vocabulary words and metacognitive awareness of word categorization and generalized usage. This application will facilitate the specialization and individualization of vocabulary and grammar lessons by allowing teachers to create customized vocabulary lists from words already in the system or with new words. If a teacher is taking her class on a field trip to the local Aquarium, for example, she will be able to create lessons on the fish the children will see at the museum. A parent could prepare lessons on the child's relatives, her schoolmates, and teachers. Most importantly, lessons can easily be easily created for the child's most recent interest.

We found that the Timo Vocabulary is effective in teaching vocabulary to English Language Learners. This result replicates previous studies carried out on hearing-impaired and autistic children with Baldi as the animated conversational tutor. In other experiments, we have also observed that Baldi's unique characteristics allow a novel approach to training speech production to both children with hearing loss [8] and adults learning a new language [9]. We look forward to new studies that continue to test the value of computer-animated tutors for language learning and related applications in education and human-machine interaction.

Acknowledgements

Baldi® is a trademark of Dominic W. Massaro. The research and writing of the paper were supported by the National Science Foun-dation (Grant No. CDA-9726363, Grant No. BCS-9905176, Grant No. IIS-0086107), Public Health Service (Grant No.

PHS R01 DC00236), a Cure Autism Now Foundation Innovative Technology Award, the National Alliance for Autism Research, and the University of California, Santa Cruz. The author thanks Roger Kimbrough for testing the children and Jennifer Schmida and David Payne for supporting this research project.

References

1. Massaro, D.W. (1998). Perceiving Talking Faces: From Speech Perception to a Behavioral Principle. MIT Press: Cambridge, MA
2. Granström B., House D. & Beskow J. (2002). Speech and gestures for talking faces inconversational dialogue systems. In B. Granstrom, D. House, & I. Karlsson (Eds.), Multilmodality in language and speech systems (pp. 209-241). The Netherlands: Kluwer Academic Publishers.
3. Massaro, D.W., Ouni, S., Cohen, M.M., & Clark, R. (2005). A Multilingual Embodied Conversational Agent. In R.H. Sprague (Ed.), Proceedings of 38th Annual Hawaii International Conference on System Sciences, (HICCS'05) (CD-ROM,10 pages). Los Alimitos, CA: IEEE Computer Society Press.
4. Massaro, D.W. (2004). Symbiotic Value of an Embodied Agent in Language Learning. In R.H. Sprague, Jr.(Ed.), Proceedings of 37th Annual Hawaii International Conference on System Sciences,(HICCS'04) (CD-ROM,10 pages). Los Alimitos, CA: IEEE Computer Society Press. Best paper in Emerging Technologies.
5. Beck, I. L., McKeown, M. G., & Kucan, L. (2002). Bringing words to life: Robust Vocabulary Instruction. New York: The Guilford Press.
6. Massaro, D.W., & Light, J. (2004). Improving the vocabulary of children with hearing loss. Volta Review, 104(3), 141-174.
7. Bosseler, A. & Massaro, D.W. (2003). Development and Evaluation of a Computer-Animated Tutor for Vocabulary and Language Learning for Children with Autism. Journal of Autism and Developmental Disorders, 33(6), 653-672.
8. Massaro, D.W., & Light, J. (2004). Using visible speech for training perception and production of speech for hard of hearing individuals. Journal of Speech, Language, and Hearing Research, 47(2), 304-320.
9. Massaro, D.W., & Light (2003, September). Read my tongue movements: bimodal learning to perceive and produce non-native speech /r/ and /l/. Eurospeech 2003-Switzerland (Interspeech). 8th European Conference on Speech Communication and Technology, Geneva, Switzerland.

Web Design for Dyslexics: Accessibility of Arabic Content

Areej Al-Wabil, Panayiotis Zaphiris, and Stephanie Wilson

Centre for Human-Computer Interaction Design
City University, London EC1V 0HB, UK
a.al-wabil@city.ac.uk,
{zaphiri, steph}@soi.city.ac.uk

Abstract. This paper reports results of a workshop on the design of electronic content for users with Specific Learning Difficulties (SpLD), particularly Arabic dyslexics. First we shed some light on the nature of the Arabic language and discuss features that account for the unique needs of Arabic users with reading disorders. Then we present recommendations for accessible web design for Arabic content in light of existing guidelines on web design for dyslexic users.

1 Introduction

The use of electronic materials is increasing and more educational content is being made available to users online. Web design guidelines promoting accessibility for users with Specific Learning Difficulties (SpLD), dyslexics in particular, have been developed [1,2]. Until recently, our understanding of dyslexia was largely confined to research on the English language. For this reason, existing accessibility guidelines fail to address some issues related to non-Latin orthographies. For accessibility issues related to presentation and content, it is important to consider the linguistic properties of the content, especially for the case of users with SpLDs. Evidence in recent studies has shown that linguistic properties of different languages such as phonological, morphological, and orthographic characteristics influence the learning of reading and writing of that particular language [3]. A few studies have begun to appear that look into characteristics of dyslexia in Arabic [4,5,6]. Design guidelines for Arabic content need to consider features of the Arabic language that account for the unique needs of users with reading disorders.

Arabic is widely spoken with an estimated 200 million native speakers. It is also the language of the Quran and is believed to be the second most widely used script in the world [7]. The spread of Islam prompted the incorporation of other cultures outside the Arab world and resulted in what is referred to as the "Arab scripted world", which includes contemporary languages such as: Farsi, Jawi, Kurdish, Pashto, Sindhi, Urdu, Dari, and Uyghuri [8].

Issues related to the design of Arabic electronic content were explored in a workshop investigating accessibility of Arabic content for users with SpLDs, particularly dyslexics. The objectives of the workshop were to identify the most

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 817 – 822, 2006.

commonly noted difficulties which dyslexics experience when reading and writing Arabic script, understand the needs of dyslexics for accessing Arabic electronic content, and present recommendations for accessible design specific to Arabic content.

1.1 Arabic Orthography

Arabic is an alphabetic language, written from right to left in cursive form, with 28 letters. Each letter has a different form depending on its position in the word (initial, medial, final). Arabic script includes long vowels, but no short vowels. Short vowels are indicated by diacritical marks, small diagonal strokes above or below characters. The Arabic orthography is either *shallow* (short vowels appear in the writing as diactritics) or *deep* (unvowelized, diacritics are omitted). Both orthographies are used for electronic content made available online. However, deep orthography dominates the majority of websites. While newspapers and textbooks utilize deep orthography aimed at skilled readers, shallow orthography is used in children's books and in formal texts such as the Quran [4]. Texts presented in the shallow orthography are easy to phonologically decode due to the consistent letter-to-sound correspondence. Very few websites provide content in this type of orthography for two salient reasons: First, a greater expansion of effort is required from the developer to include diacritics, especially for larger websites. Secondly, the linguistic skill required for the correct application of diacritics can be intimidating even to native speakers working as developers.

Arabic, like other Semitic languages is also Bi-Directional (BiDi for short). Most of the text is written from right to left, while some of the text (like numbers) is written from left to right. This requires special accommodation for editors and browsers and only recently have products been able to address this issue and facilitate bidirectional support. Arabic is also a highly homographic language where the same word can carry many different meanings [9].

1.2 Dyslexia in Arabic Readers

Very few studies have investigated dyslexia in native readers of Arabic. Common problems of dyslexia that have been observed and reported in the literature of Arabic dyslexia include (Adapted from Al-Sarhan [10] and Al-Sartawi [11]): Letter and word recognition for letters that are similar orthographically or phonetically, letter and word omissions and additions (including diacritics - short vowels), short term memory problems, letter and number reversals, and disturbances in visual and audio processing. Due to the idiosyncratic nature of dyslexia, these difficulties afflict individuals in varying forms and degrees of severity.

Studies by Abu-Rabia [12,4] have investigated dyslexia in Arabic and showed similar characteristics to those in other languages. These studies have shown that Arabic dyslexics have poor phonological processing skills and working memory (cognitive processing) compared to normal readers. A significant difference is the visual processing required for diacritics in Arabic [13]. Elbehari argues that factors such as the cursive nature of the Arabic script, inflectional morphemes, different graphemes, and discrepancy between phonology and orthography in unvowelized scripts could possibly be contributing to the manifestation of dyslexia in monolingual Arabic speakers [6].

2 Methods

A workshop on assistive technology for users with specific learning difficulties related to reading and writing of the Arabic language, particularly dyslexics, was held in Riyadh, Saudi Arabia on January 3, 2006. The objective was to review guidelines promoting accessibility of electronic content for dyslexic readers and explore issues related to Arabic content. A total of fifteen participants included specialists in specific learning difficulties and educational software development. Four participants were experts in SpLDs from the college of special education in King Saud University, four specialists in SpLDs represented the General Secretariat of Special Education, four special education teachers with experience in working with dyslexic students at Ministry of Education public and private schools, and three computer scientists with experience in developing Arabic software from Prince Sultan University and the Ministry of Education. Issues affecting the accessibility of Arabic electronic materials such as readability, user control over presentation, and information architecture were discussed.

To address the objectives of this workshop, an Arabic translation of the checklist for web design, for Dyslexic and Visually Impaired readers (DYVI), was reviewed [2]. Each item/guideline was presented with an example of an Arabic website or electronic content. For each guideline, we discussed the underlying dyslexic difficulties and their design implications. Furthermore, we asked participants to rate the significance of each guideline with relation to Arabic electronic content on a scale of one to five, one being not significant and five being very significant. Other issues, specific to Arabic script and not addressed in the published guidelines, were also raised and discussed in our review.

3 Results and Discussion

Participants were in greatest agreement about the benefits of computer-based materials that could provide content that is configured to users' needs. The capability of providing alternative methods of content presentation such as audio and video was believed to be helpful for dyslexic students. Audio, in particular, was reported from participants' practical experience (due to the availability of audio recording of some text based materials in the curriculum) to be a helpful aid for students in understanding content. It has been shown that for many users with SpLDs, "the accuracy of the perception of the spoken word is greater than the perception of the written word" [14]. It is also noted that many participants were unaware of how simple changes to the interface could affect the ability to read, as has already been demonstrated in studies on users of the English language [15,16]. In the workshop review of the DYVI checklist, most of the items were found to apply directly to

Arabic content and were rated very significant with the exception of a few. For example: visual contrast, minimum font size of 12pt, and avoiding light text on dark background had the least average rating among participants. It was also noted that some items, in the DYVI checklist, do not apply to Arabic content such avoiding capitals mid-line and using full stops for abbreviations and acronyms.

It was agreed that alignment of text should be from right to left due to the nature of the script. However, fully justified text was not seen as a problem because in Arabic, words can be stretched to fill the width of the line instead of introducing inconsistent spaces between words as in English. This type of justified text was reported by participants, from practical experience, as being quite helpful in letter recognition and associated diacritics, as well as providing a visual aid when reading from right to left and moving from one line to the next.

As for font types, participants rated the clarity of scripts that have flowing cursive form with full and deep curves and straight and vertical uprights such as Naskh much higher than angular types such as in Koufi and Andalus types. The preferred text font types, rated by participants, were Arabic Transparent followed by Simplified Arabic Fixed. As for text size, the preferred size was in the 16-20pt range. This is not surprising as Arabic script is usually presented in larger sizes than English, both in print and electronic form. In terms of emphasizing text, results were in agreement with the recommendation of using bold rather than italic or underline. Many suggested the use of color or a combination of both bold and color as is used in many Arabic textbooks for early elementary levels. This is also in agreement with the fact that "Adding color can increase the likelihood that the information will enter the long term memory" [17].

The recommendation of using short sentences was given an average 4.7/5 significance rating by participants. It was agreed that shorter sentences are easier to read and comprehend. However, adhering to this guideline might present a challenge, as the nature of the Arabic language allows for longer sentences that span many lines, a distinct contrast to English.

Spacing between letters and words is a problem in the presentation of Arabic texts. Six letters in the Arabic alphabet cannot be joined, and consequently create spaces between letters within words in Arabic. Readers must learn to distinguish word boundaries (space between words) from breaks in the cursive script within a word. This might add to the difficulties dyslexics face in the visual processing of text. It is recommended that spacing between words be increased.

Perhaps the most interesting point which arose during the review of the guidelines was the significance of diacritization (vowelization) when promoting accessibility of Arabic content. Participants agreed that the use of diacritics could improve reading accuracy and comprehension, especially for beginning readers. In a study on the role of vowels in the reading of Semitic scripts, Abu Rabia [12] states that "If texts are not vowelized, these short vowels must be deduced by the reader, which is a cognitively demanding task". Results in another study [9] indicated that vowels had a significant effect on reading accuracy of poor and skilled readers. This is not an issue in English, as vowels are present in all text. For example, consider the English words "heard", "herd", "hard", and "hired". A deep, Arabic orthography would reduce these words to "hrd" and then leave it to the individual reader to decipher the word and add its missing vowels depending on the context. Furthermore, the extensive use of deep orthography in Arabic electronic content adds another barrier to accessing its content to both visually impaired and dyslexic readers. The visually-impaired group is affected because screen readers are prone to mispronounce words that do not have diacritics, thus rendering content incomprehensible. The dyslexic group is also adversely affected by being forced to read without diacritics, because without them they must conjure a combination of contextualized cognition and visual memory, thus adding to their cognitive burden. While diacritics help to disambiguate words, they also have the negative effect of adding to the visual complexity of the script. Reading with diacritics requires high levels of visual discrimination and memory skills. This in turn reduces the reading comfort of both beginning and experienced readers. A suggestion was made to offer text in both formats in electronic form and give the user the option to view or hide the short vowels. Offering levels of vowelization, full or partial, would address the visual complexity of the script. Some assistive technologies have been developed to automatically add diacritics, such as Sakhr's Automatic Diacritizer [18]. However, this solution is specific to a limited range of commercial products and not affordable to most users. The Unicode standard accommodates the representation of diacritics. It is recommended to consider automatic diacritization of text in applications so that users can control the appearance of diacritics in electronic Arabic content.

5 Conclusion

This paper presents recommendations for promoting accessibility of Arabic content derived from practitioners' experience. Suggestions were made to offer diacritization to facilitate comprehension and to consider simple and clear Arabic typography in presenting text to increase readability. These recommendations need to be further supported by usability research and tested experimentally at a later stage with dyslexic users of Arabic content. It is important to consider the unique features of the Arabic language and orthography and their corresponding design implications in promoting accessibility of Arabic electronic content. Although this research is in its early stages, we feel that our initial findings can give some guidance to practitioners on how to provide accessible Arabic content. Recognizing the diversity of individual differences between dyslexics when providing electronic content would have a substantial benefit for dyslexics and consequently for a much wider section of society. In the Arab world today, there is a great lack of awareness about Specific Learning Difficulties in general, and dyslexia in particular. More research in this area is required to better understand the needs of the dyslexic population so that they can acquire literacy skills and achieve their full potential.

References

1. BDA - British Dyslexia Association (2004). Dyslexia Style Guide. Available online at: `http://www.bdadyslexia.org.uk/extra352.html` [date of access 13.01.06]
2. Evett, L., Brown, D. (2005). Text formats and web design for visually impaired and dyslexic readers – Clear text for all. Interacting with computers, 17, 453-472

3. Goulandris, N. (2003). Dyslexia in Different Languages. Cross-Linguistic Comparisons. London, UK: Whurr Publishers
4. Abu-Rabia S. (2004). Dyslexia in Arabic. In Smyth, I., Everatt, J. Salter, R. (Eds.), International Book of Dyslexia: A Cross-Language Comparison and Practice Guide (pp. 31-38). West Sussex: John Wiley & Sons
5. Beland, B., Mimouni, Z. (2001). Deep Dyslexia in the two languages of an Arabic/French Bilingual Patient. Cognition, 82, 77-126
6. Elbehari, G. (2004). Dyslexia in Egypt. In Smyth, I., Everatt, J. Salter, R. (Ed.), International Book of Dyslexia: A Guide to Practice and Resources (pp. 79-85). West Sussex: John Wiley & Sons
7. Holes, C. (1995). Modern Arabic: Structures, Functions and Varieties. New York: Longman Group Ltd
8. Milo, T. (2002). Authentic Arabic: a Case Study. 20th International Unicode Conference. Washington, DC. Available online at: `http://www.tradigital.de/specials/casestudies.htm` [date of access 27.12.05]
9. Abu-Rabia, S. (1998). Reading Arabic Texts: Effects of Text type, Reader Type, and Vowelization. Reading and Writing: An Interdisciplinary Journal, 10, 106-119
10. Al-Sarhan, M. (2001). Report on Students with Specific Learning Difficulties. Ministry of Education (in Arabic). `http://www.riyadhedu.gov.sa/alan/fntok/special/soobat.doc` [date of access 27.12.05]
11. Al-Sartawi, Z., Khashan, A., Al-Sartawi, A., Abujouda, W. (2001) Introduction to Specific Learning Difficulties. (in Arabic) Riyadh, Saudi Arabia: Special Education Academy Publications
12. Abu-Rabia, S. (2001). The Role of Vowels in Reading Semitic Scripts: Data from Arabic and Hebrew. Reading and Writing: An Interdisciplinary Journal, 14, 39-59
13. Abu-Rabia S., Share, D., and Mansour, M. (2003). Word Recognition and Basic Cognitive Processes among Reading-Disabled and Normal Readers in Arabic. Reading and Writing: An Interdisciplinary Journal, 16, 423-442
14. Evans, D., Blenkhorn, P. (2005). Designing a Screen Reader for Fully Sighted Users. HCI International 2005, Las Vegas, Nevada
15. Gregor, P., Dickinson, A., Macaffer, A., Andreasen, P. (2003). SeeWord – a Personal Word Processing Environment for Dyslexic Computer Users. British Journal of Educational Technology, Vol 34 No 3, 341-355
16. Kurniawan, S., Conroy, G. (2006). Comparing Comprehension Speed and Accuracy of Online Information in Students with and without Dyslexia. In Kurniawan, S. and Zaphiris, P. (Eds.), Advances in Universal Web Design and Evaluation: Research, Trends and Opportunities. Hershey: Idea Group Inc
17. Rainger, P. (2003). A Dyslexic Perspective on e-Content Accessibility. Available online at: `http://www.techdis.ac.uk/seven/papers/dyslexia.html` [date of access 27.12.05]
18. Sakhr Software. `http://www.sakhr.com/Sakhr_e/Technology/Diacritization.htm`

Flexibility in Virtual Environments – A Fully Adjustable Virtual Classroom

Zoltán Geiszt, Cecília Sik-Lányi, and Péter Károlyi

Department of Image Processing and Neourocumputing, Pannon University
8200 Veszprém, Egyetem u. 10. Hungary
geiszt@vision.vein.hu, lanyi@almos.vein.hu,
chartless@freemail.hu
http://www.knt.vein.hu

Abstract. The fear of public speaking is one of the most wide-spread social problems of the world today. It is present in early childhood as well as in adultness, and causes problems with further influence on the child's entire life. The hesitation or the inability to speak in a heart-quaked situation can easily lead to bullying from classmates or even the teacher can misinterpret it as lack of knowledge or unwillingness to give the answer at all. For helping such children overcome their fear we developed a virtual environment imitating a classroom with its usual elements. The user immerses the virtual environment via a head mounted display where he or she must solve tasks in a "living" classroom of noisy, commentating classmates and a virtual teacher. A teacher or therapist supervises the acting in the virtual class and reacts to the doings of the user through speaking to him directly or allocating a speech utterance or other act to one of the virtual classmates. Due to its complex nature and complete adaptability the virtual classroom proved to be a very useful tool in helping such children and it is already in use in a primary school.

1 Introduction

Many people in the world today have some innate fear. We fear of little insects, or great depths, and often of the opinion of others on us. This latter one reaches its peak in cases of public speaking, when stage-fright got us. The effect is not present only in adulthood, but also causes serious problems in years in primary education. The hesitation or the inability to speak in a heart-quaked situation can easily lead to bullying from class-mates or even the teacher can misinterpret it as lack of knowledge or unwillingness to give the answer at all. Through these implications the problem is further deepened and poses a very negative influence on the entire life of the concerned person.

This problem is obviously one from the branch of problems which is the best to fight through practice. But everyone is reluctant to get themselves involved in unpleasant situations - this is the real knot in the problem. But virtual reality offers a new manner for building and taking part in such situations through its multimodal interactivity [7]. It eliminates the most of the possible negative effects of a real

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 823 – 830, 2006.

situation, but keeps most of the positive ones. Even it proved trustworthy before in a similar field of use: Riva and his colleagues used it in treatment of phobias [4].

It is a fascinating instrument giving a wide variety of tools and even can work without supervision via pre-defined automatism. A pre-programmed world-model in a virtual environment (VE) only follows standards based on some kind of generalisation and targets a generalised problem. But an individually adjusted system targets the very problem of the concerned individual person, presents his or her environment, surroundings and acquaintances.

The spark for the idea of creating a virtual classroom to overcome the fear of public speaking came from the Integrated Media Systems Center's Virtual Classroom Research [1] and the work of Rizzo and his colleagues. They created a general classroom for measurement of the distracting factors during class-work [5,6], we wanted another one for treating anxiety disorders. From the beginning we knew that our virtual class should be highly flexible and adjustable to meet any needs and therefore conducted the development process according to this premise. The outcome of our project has become such a virtual class in which nearly everything can be modified and replaced: the floor-plan of the room along with the number and position of the inner objects and the living residents. Even the textures can be replaced, through which feature the very same class from real life can be imitated in the virtual reality. Just imagine a virtual place with your familiar surroundings and people – this is the Adjustable Virtual Classroom (Fig. 1).

Fig. 1. One from the uncountable variations of the virtual classroom

2 The Development Process

Development process of virtual reality related applications is usually divided into two parts. One is the modelling of the elements, objects, agents of the virtual world, while the other is the creation, the programming of the engine animating the models, bringing the whole environment into life.

In our case the modelling of the objects was carried out via using the 3D designer Maya 6.5, while the modelling of virtual agents was done by a specialised designer, Poser 6. The engine itself was written in Visual C++ 6.0 with DirectX 9.0 API (Application Programming Interface).

For the refinement of textures, mostly taken from photographs another two programs, UVMapper Classic and the well-known Adobe Photoshop CS2 were used. To have the models transformed into DirectX format Panda DirectX Exporter and 3D Studio Max 7.0 gave a hand to us. Visual C++ 6.0 was accompanied by the MFC (Microsoft Foundation Classes) in the creation of the Editor of the program, in which anyone can create their own virtual classroom using the functions of a well-designed interface. It also serves as the editor of the tasks that the user has to solve when immersed in the virtual environment. For the storage of these tasks along with actions, animations and other data of the virtual agents, we made good use of XML (Extensible Markup Language) [2,3].

3 The Use of the Adjustable Virtual Class

For the Adjustable Virtual Class (AVC) is a photo-realistic, interactive imitation of a primary school classroom, the following requirements are understandable.

The necessary hardware configuration is the following: an IBM PC compatible computer, 1GHz or larger processor, 256 MB RAM, 50 MB free space on hard disk, DirectX 8.0 compatible video card with 128MB memory, DirectX compatible soundcard, and a mouse. The head mounted display is not compulsory (only suggested), a normal monitor will do.

The necessary software configuration is Microsoft Windows 98 or Microsoft Windows 2000 or Microsoft Windows XP operation system, and DirectX 9.0 or fresher.

About the installation of the AVC there is no much to be said. If the software does not start automatically, the user can start the installation by running the Setup.exe file. The two main sections of the software, the Editor and the Viewer are granted two separate icons and entry in their shared Windows Start Menu folder.

Let's see the wide array of adjustable features in the Editor first. The two most important menus are the Options menu and the main menu containing the submenus for the setting up of the classified features. In the Options menu the usual entries can be found: screen resolution, colour depth, etc., and the language of the Editor interface. Currently it is able to communicate in both Hungarian and English.

Under the People submenu of the main menu a list of features concerning the avatars can be found. The model itself and the textures can be imported from files, as well as the audio material for a complete of five speech utterances. The suggested picture format is JPEG, while for the audio it is WAW.

The Exercises dialogue window is responsible for the setup of the task to be solved by the user while acting in the virtual environment. Both the question and the answer can be given a text and a picture later appearing on the blackboard of the virtual classroom, even an audio material can be imported and associated to them. The limit on the text of the question and the answer is limited to 200 characters.

The Class dialogue serves for the setting up of the classroom community. The floor-plan itself can be given here as well as the number of the pupils along with their whole representation (model with textures and voice).

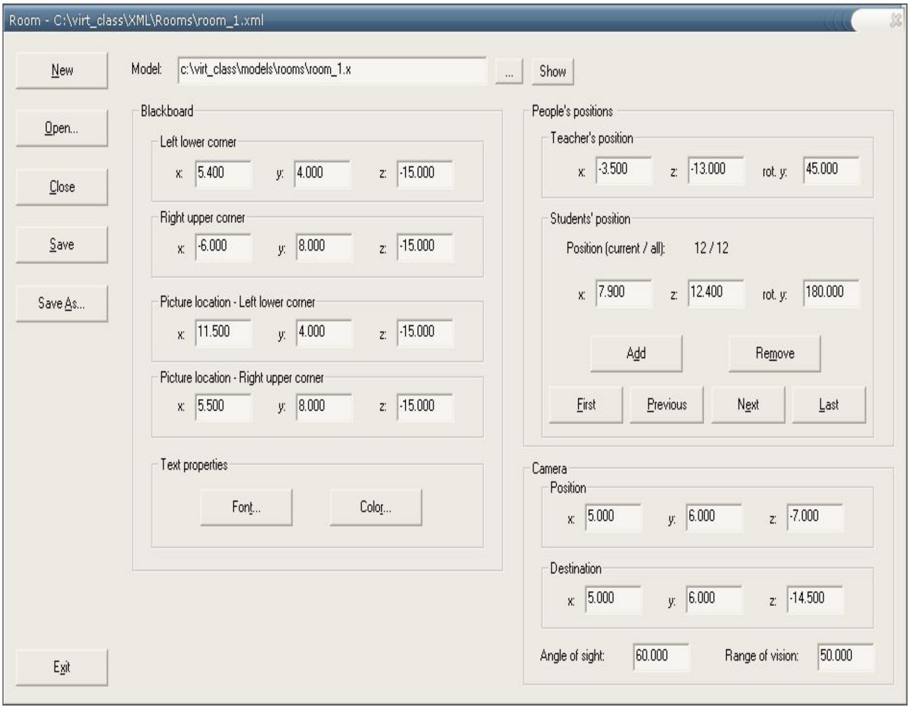


Fig. 2. The parameters of the classroom

The Room submenu (Fig. 2) contains the entries concerning the spatial structure of the room itself. A pre-made model can be selected for the room, in which the coordinates of the main objects must also be given. One of these objects is the blackboard which gained its importance through the fact that the position of the texts of the questions and the answers and their pictures depend on it. The position of the teacher and the pupils must also be given, with taking attention to their order since this determines their identifier number used in the Class dialogue. Lastly the initial position and direction of the camera is also set up here.

The task of the Viewer software is to display the virtual classroom whose settings were determined earlier in the Editor (Fig. 3 & 4). While the user is immersed in the virtual environment with the use of a Head Mounted Display a trainee can follow his or her acting in the virtual class through the monitor and can give instructions, comments, feedback. In addition the teacher is able to control the virtual environment as it was mentioned before. Numerous shortcuts are offered by the software or the control of the environment.

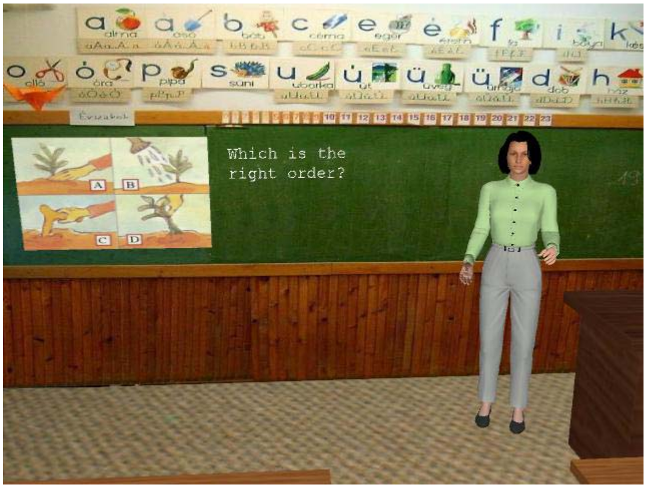


Fig. 3. The virtual teacher and a task on the blackboard

Fig. 4. The virtual pupils

For example the function keys from F1 to F10 select the virtual pupil with the correspondent number, while F12 stands for the virtual teacher. If an avatar is selected, with pressing of numbers 1 to 5 an activity, set before in the Editor, is associated to him or her and its performance starts immediately.

Another branch of keys are related to the handling of tasks. Numeric + and – shows or hides an exercise, while right and left cursor keys move on and back in the list of tasks and the down key selects one randomly.

The AVC can be removed from the computer by the „Removing AVC" entry in the folder of the software.

4 The Test of the Adjustable Virtual Class

The first usability test of the software was conducted in a very special field: in Flóra Kozmutza Ground School Skill-Developing Special School and College in Veszprém, Hungary. Special field for the AVC, since this institution is formed for the education of mentally disabled children. Here not only primary school teaching material is taught but the profession of weaving for the older pupils.

Our subject group (Fig. 5) contained mostly children with average level disability and a few with such sever conditions which are most likely make them unable to learn to read. The complete group measures 64 children, with a further drawback: they had had a very little amount of experience with computers as opposed to average children.

Fig. 5. Exploring the virtual classroom

Several tasks in the virtual environment were given attention during the test, the first of which was the development of attention. The teachers taking part in the assessment of the test stated that the virtual environment gained the attention of the pupils. The easy exploration of the class by the use of the HMD did not interfere with the level of attention, everyone concentrated on the tasks for the whole duration of immersion. HMD even helped the children to fix their gaze, which is very difficult for them since they are much more likely to be distracted by any event outside than the average children.

Secondly their interpretation of tasks were assessed, where the same beneficial influence of the HMD on fixation gave a huge help them in concentration exclusively on the task. All the illustrations for the tasks were of good quality and ostensive. The only drawback mentioned by the test-supervising teachers is the fact that the tasks were given only in written form and they had to read it aloud for the pupils. Our first amendment to the software is going to be the filling of this gap.

Not only the interpretation of tasks were focused on, but their solving as well. The teachers' opinion was also very positive on this topic, they agreed in the fact that the users could already feel themselves an active part of the virtual environment. All of the children were very active in the task-solving and very cooperative. But their independence and activity in the classroom is the main concern of the project, and there were a few suggestions to further amend it. First the use of the mouse cursor to select the correct answer. Then the expansion of the slightly delayed feedback of the virtual teacher on the rightfulness of the answer with a direct and immediate comment from the VE. This is needed especially in the case of children with more severe level of disability. Its embodiment could be an audio or visual element or as best the combination of both, but must be as plain as possible.

The last field of the test measurements were the one concerning feedback. In the previous paragraph we have already stated that the children were very cooperative in the solving of tasks, which was likely a consequence of the activity of the virtual teacher. She commented the users' answers with speech and gestures, which was so exciting for them that they were looking forward to it when they were doing the given tasks. In spite of this fact we would expand the software to allow the option of an instant direct feedback, which was laid claim for in the previous paragraph. The other recommendation called for the possibility of retaking a task in cooperation with the teacher if it was failed before.

Fig. 6. Cheerful experience

5 Summary

Our project in the field of treating the fear of public speaking in primary school is coming to and end. The first test proved the usability of the program well, the students found such a great fun in its use that they insisted on its continuous use in the future.

The participating teachers were nearly as amazed as the children and they stated unanimously that the software is an immensurable help in their work, a peerless appliance.

AVC is not a rigid construction, it is as flexible as it was intended to be, not only in terms of its programmed features but also in the willingness of its programmers to amend it. This way the implementation of the mentioned suggestions has been started already, and will be finished soon. Our virtual class should be really adjustable in every sense of the word.

After the test in Flóra Kozmutza School, the software is brought to various other primary schools with average pupils, and its testing is going to start with the beginning of the new semester. Due to its already gained achievements all the cooperating teacher of the project are looking forward to see it in work with their students. Adjustable Virtual Class is not only a software package for treating the fear of public speaking but also a fascinating source of joy for children of mental disabilities (Fig. 6).

References

1. Integrated Media Systems Center – Virtual Classroom Research `http://imsc.usc.edu/research/project/virtclass/`
2. Jurecka, J.: Working with the DirectX .X File Format and Animation in DirectX 9.0. GameDev.net (2003)
3. Nyisztor, K.: Graphics and Programming of Games with DirectX (Grafika és játék-programozás DirectX-szel), SZAK Kiadó Kft., Budapest, Hungary (2005)
4. Riva, G., Wiederhold, B. K., Molinari, E.: Virtual Environments in Clinical Psychology and Neuroscience. IOS Press, Amsterdam, Netherlands (1998)
5. Rizzo, A. A., Bowerly Y, T., Buckwalter, J. G., Schultheis, M., Matheis, R., Shahabi, C., Kim, L., Sharifzadeh, M.: Virtual environments for the assessment of attention and memory processes: the virtual classroom and office. Proc. 4th Intl Conf. Disability, Virtual Reality & Assoc. Tech., Veszprém, Hungary (2002) 3 - 6
6. Rizzo, A. A., Bowerly Y, T., Shahabi, C., Buckwalter, J. G., Klimchuk, D., Mitura, R.: Diagnosing Attention Disorders in a Virtual Classroom. IEEE Computer Vol. 37 (2004) 87-89
7. The Encyclopedia of Virtual Environments. `http://www.hitl.washington.edu/scivw/EVE/`

Temporal Orientation Panel for Special Education

Jorge L. Falcó[1], Carmen Muro[2], Inmaculada Plaza[3], and Armando Roy[1]

[1] Grupo TecnoDiscap – Dpto. Ingeniería Electrónica y Comunicaciones,
Centro Politécnico Superior (CPS) – Universidad de Zaragoza
María de Luna, 1, 50018, Zaragoza, Spain
{jfalco, armanroy}@unizar.es
[2] Grupo TecnoDiscap – Dpto. Fisiatría y Enfermería
Enfermería – Universidad de Zaragoza
C/Domingo Miral s/n, 50009, Zaragoza, Spain
cmuro@unizar.es
[3] Grupo TecnoDiscap – Dpto. Ingeniería Electrónica y Comunicaciones,
Escuela Universitaria Politécnica de Teruel – Universidad de Zaragoza
Ciudad Escolar s/n, 40003, Teruel, Spain
iplaza@unizar.es

Abstract. In the present work an electronic panel designed for helping people with deficiencies in their understanding of the sense of time is showed. The device will be installed in special education classrooms. It tries to meet the following objectives: 1) To provide disabled people with temporal orientation and to help in learning the use of conventional clocks. 2) To make the concept of temporal grouping easier. The grouping mechanism reflects the ability to use the information about task or work at some point correlated with temporal indicators. 3) To make anticipation of the sequence of events possible for autistic children. The results of the device's evaluation will allow authors to improve the design. Thus, it could be possible to extrapolate the use of the device on homes for disabled children or for elderly people with cognitive disabilities or even in the early stages of senile dementia.

1 Introduction

The *"Technical aids for disabled persons"* can be defined as a product, instrument, device or technical system used for people with disabilities. They have been specially designed, manufactured, supplied or commercialized in order to prevent, compensate, mitigate or neutralize any lack or disability. The ISO 9999:2002 provides the standard definition [1]. The technical aids are usually named as "assistance devices" or "support technology".

Among the different kinds of aids, we can find the aids destined for the training of capacities and skills (number 05 in the international standard ISO 9999:2002).

Clocks are timepieces that show the time of day. They can be considered as aids for training in lack of understanding of the sense of time (05 15 09 ISO 9999:2002 code). Also the can be classified as an aid for personal care and protection (09 51 ISO 9999:2002 code).

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 831–838, 2006.

In the present work, the authors show the design of a temporary orientation panel developed in order to meet the following objectives: 1) To provide disabled people with temporal orientation and to help in learning the use of conventional clocks. 2) To make the concept of temporal grouping easier. The grouping mechanism reflects the ability to use the information about task or work at some point correlated with temporal indicators. 3) To make anticipation of the sequence of events possible for autistic children.

The paper is divided into several sections: the origin of the panel (the state of the art and the initial hypothesis), its physical implementation and the guidelines used in order to evaluate its operation in two special education classrooms. Finally, future lines of work are explained.

2 Initial Hypothesis

It is necessary to explore the "state of the art" in order to understand the initial hypothesis of this work. Several organizations have developed devices for temporal training and re-education for disabled people. From the analysis of their works it is possible to find several kinds of devices or systems. For example, the clock developed for CERTER [2] or the device of Handitek [3].

1. CERTEC (Center for Rehabilitation Engineering Research Lund Institute of Tech-nology) [2] is a Swedish institution. It works in aid systems for people with disabilities. From the CERTEC's studies it can be concluded that the quarter of an hour is the simplest temporal unit to assimilate the sense of the time for disabled people.

   The proposed solutions are clocks using color, shape, scent or patterns.

   One example consists of a row of lights, a system for measuring time based in the full/empty concepts. Each light corresponds to a quarter of an hour, i.e. 15 minutes. The lights switched off as fast as the time pass. Thus, the day starts with the full concept (all lights on) and the journey finishes with the concept of empty (lights off).

   The basic idea is to simplify the measure of the time. Most clocks demand logical-mathematical operations and an ability to make judgments, all of which can be problematic for dissabled people. To measure time in length doesn´t place such high demands on the user´s cognitive abilities.

   The prototipe designed for CERTEC consists of three blocks of LED´s representing the morning, the afternoon/evening and the night. In the casing of the device there are several drawings to illustrate the task that should be carried out at some point.
2. HANDITEK (a Swedish company also) is a technology development company in the disability field. Its mission is to incresase the independence of people with cognitive disabilities by developing assistive technology that meets their needs. To fulfill its mission, it develops assistive devices; conduct studies and offers training programs [3].

   This company, with The Swedish Handicap Institute, implemented a portatil clock: the Quarter Hour Watch (QHW). This device is a time aid for people with intellectual disabilities. The QHW does not show the actual time like ordinary clocks. Instead the time remaining until an event starts, is graphically displayed by dots (or, in the timer mode, the time remaining until an event ends). Each dot corresponds to a quarter of an hour.

To see how much time is left until an event starts, a picture card is placed in the QHW. On the front of the card there is a picture that symbolizes the event. On the rear or the card the start time is programmed. Programming the start time is simple. It is punched out on a disposable clock face made of plastic, which is clipped on the rear or the picture card. When the time arrives, the QHW beeps.

An initial hypothesis follows as a summary of previous works: *"the quarter of an hour is the simplest temporal unit to assimilate the sense of the time for disabled people. Thus, the right solution is to work with the "Quarter Hour Watch" a system for measuring time based in full/empty concepts".*

Three case studies on the use of the Quarter Hour Watch are written by two occupational therapists (Rose-Marie Remvall and Karin Månsson). This study is based on their extensive experience of introducing this special watch in the early nineties. They used the Quarter Hour Watch performing the various daily tasks [3].

After the initial phase of adaptation, the results show that people has learned to use the indicators on the display in order to determine the time remaining before an event is to take place. Also they used pictures for various activities (at the day centre and at home).

Moreover their perception of time shows a substantial increase.

Thus, it is possible to conclude that the Quarter Hour Watch help disabled people to establish a temporal reference and to learn the measurement of the time.

But, there are still *several situations to be considered*:

– To gain the attention of the customers and to warn in advance: it is necessary try to gain the attention of the disabled person and to inform him about changes that will happen.
– Flexibility: in order to facilitate professionals to programming temporal indicators.
– Globalization of the Environment: to integrate information about the environment by means of the programming of specific tasks.
– Information transmission: in order to facilitate understanding of information.
– Use of melodies and way of lighting: designed to improve comprehension and the capacity of gain the attention of the customer.

These needs mark the *specific objectives* of this work. In particular, it is essential to try to obtain a tool that allows disabled people:

– To perceive the pass of the time in a simple way and to know the moment of the day. It is expected to provide a support to people who doesn´t understand the traditional clock and to increase independence in the planning of their time and in performing the various daily tasks.

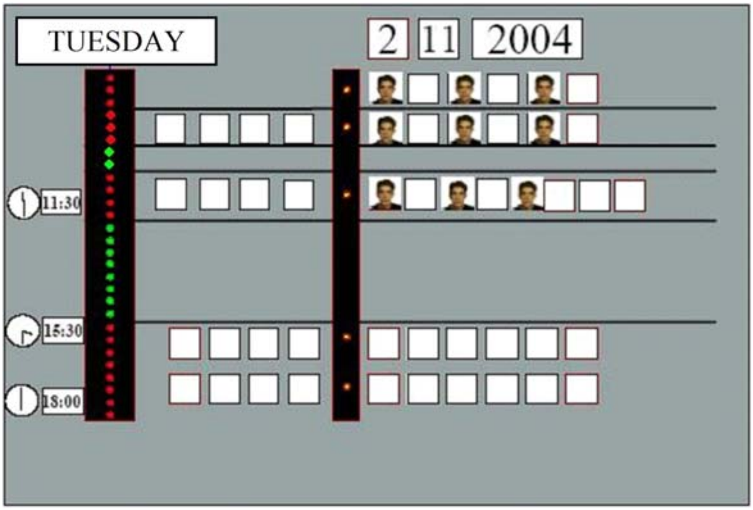


Fig. 1. Temporal Orientation Panel

- To inform in advance about the changes and situations related with temporal indicators: Therefore users can feel more in control of the planning of their time, and thus feel calmer. This point is especially important for people with intellectual disability with autistic traits.
- To train for the planning of tasks. This objective (related with the first) con-tributes to increase of the independence.

3 Description of the Panel

The panel was designed with the help of teachers of Alborada. Alborada is a Special Education State School located in Zaragoza (Spain) [4]. Its students are children with learn disabilities and youngsters with autistic traits. The final prototype will be used in classroom (see Fig. 1).

This is a grey 2x1,5 metres panel, with satin texture. It allows to write and to rub off using felt-tips for whiteboard. It enables also to stick and to unstick cards. These cards are similar to the drawings used in clasrooms to represent the timetable.

Four areas can be distinguished:

1. *Left: temporal representation.* Is made up of a row of 28 lights (LEDs) in an upright position. They turn off every quarter of an hour and represent all the school day.

   In this preliminary version we have chosen the red (represents the time outside the clasroom) and green (time in the clasroom) colours.

   These colours could be modified in future versions. In the left side, we save a place in order to allow teachers to show the standard hours (number or traditional clock).

2. *Up: day and date.* This area allows teachers to work on the day of the week and date.

   At the beginning of this work authors considered the possibility of representing a wall calendar: month and season. Nevertheless the panel was already crowded and this option was rejected.
3. *Centre – left: timetable for common activities.* There are separate cards in a horizontal position. Every line symbolizes one of the activities to be done in the classroom. The cards permit a graphic description of the activities. They also refer to developed tasks. In this way, the cognitive jump is minimal because all the children have a common temporal reference.
4. *Right: activities for a specific student.* Finally, this area is important as an individual congnitive support and in order to provide autistic students with personal information in advance. The objective is to reduce their ansiety, blockage or stress regarding time.

   The area has an orange luminous element. This light starts to flicker when the expected time has come. A card is related with this warning. The card shows the photo of the student who should do the activity (to go to a workshop, to go to the therapist, a tutorial session, etcetera).

   Thus, the activity intervals are indicated and the disabled student is advised about the task to do.

The panel has an associated PC in the classroom. This computer is the element in charge of generating melodies in order to identify the music depending on a previously designed strategy.

For instance:

- common changes
- to give information in advance
- at the moment to carry out an action
- to specify the student who should do some activity.

The panel can be connected to the PC. In this way, to turn luminous elements on is easily programmable (see Fig. 2). The global system is very flexible and easy to use.

Teachers can also specify the hours from the PC and control the LED´s brightness depending on the classroom light.

The *architecture of the hardware* consists of three big boards:

- The central controller: connected to the PC with a wireless serial communication. It allows the re-synchronization and daily programming. It also manages a serial communication with Serial/Parallel Registers that turn LEDs on. With this communication, the system controls the duty cycle (therefore, the intensity of the light). In this way, it is possible to choose the LED that will turn on, updating the symbolized time, the flickering, etcetera. The central controller organizes the LEDs corresponding to changes in scholar or daily activities. Another function is to generate melodies in order to facilitate the stand-alone operation once it has been programmed with the PC.

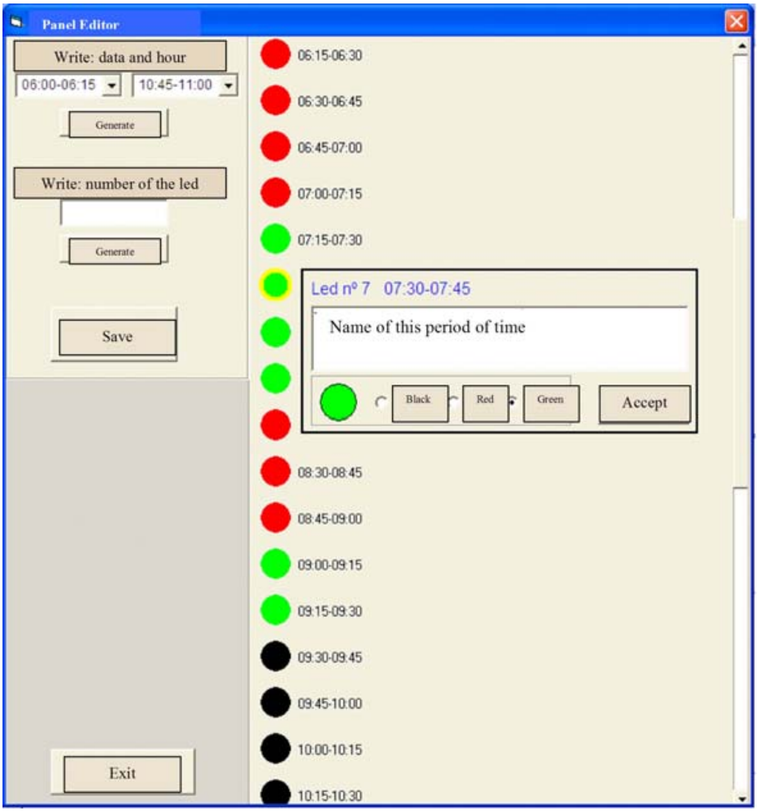


Fig. 2. Panel editor

- Three boards have the Serial/Parallel Registers that turn LEDs in the temporal representation area (on the left of the panel). The boards are strung together in order to facilitate assembly. Serial data, brightness data and Power Supply use this kind of communication.
- Other boards turn on/ switch off elements in the right area.

Before finishing this section, it is necessary to underline the current regulation that applies to the temporal orientation panel. In the case of use in a school, it is com-pulsory to test it according to the International Standards IEC 61000-6-1 [5] e IEC 61000-6-3 [6] (about Electromagnetic Compatibility) and according to the IEC 6959-1 [7] standard.

4 Assessment

The first tests were developed in a classroom to prove the correct visualization of the luminous elements depending on the natural lighting (sun, night, rainy day and etcetera). The results indicated the need of control the intensity of the luminous elements.

The present prototype will be tested in the classroom of a state school from January 2006 to June 2006. The choice of this scene has been motivated by several criteria:

- Children: The use in special education schools will allow evaluating its usefulness to train and educate in temporal concept and assessing its use with autistic children.
- Professionals: In this kind of schools, there are a lot of professionals specialised in personal independence. They have showed their sympathy interest in this project. Authors consider their conseil as an inestimable help.

The following hypothesis will be evaluated:

- The system improves temporal orientation of especial education students with autistic traits or intellectual disabilities who cannot manage a conventional clock in a satisfactory way.
- The temporal orientation panel reduces or even removes blockage regarding the concept of time exhibited by autistic children.
- Special education students improve their capacity to manage time and to plan activities.

In order to support the assessment, the "quality model" concept will be used. The International Standard ISO/IEC 9126-1:2001 [8] describes this concept. Also the quantitative evaluation plan template of the International Standard ISO/IEC 14598-2:2000 [9]. Although, these standards were written for software products, they are not specific to software, but are also applicable to other complex products [10].

5 Future Lines of Work

From the results obtained in the panel's assessment it will be possible to tackle other lines of work:

1. *To improve the physical implementation of the system.*
   - To clarify the suitability of different strategies in the use of melodies.
   - To analyse the colour's impact. In this way, it will be possible to deduce the more suitable colours for use in the panel.
   - To study the benefit of incorporating the module of melodies in a portable wireless device giving it to each student. The final idea will be to inform every child about his activities.
2. *Users.* In a second phase, the obtained results will be extrapolated in order to work with elderly people with cognitive disabilities or even in the early stages of senile dementia, helping combat the reduced cognitive functioning that can affect different aspects of daily living.

3. *Location.* The temporal orientation panel has been designed to use in classroom of special education school, but it will also be used at home. In this way, the system will facilitate the continuation of training and helping.

   Similarly, and related with the previous line of work, authors plan to install the device in residences for elderly men and women.

References

1. ISO 9999:2002. Technical aids for persons with disabilities. Classification and terminology.
2. http://www.certec.lth.se/doc/certecscore/. Web side of CERTEC – Center for Rehabilitation Engineering Research Lund Institute of Technology. Last visit: January 2006.
3. http://www.handitek.se/english.htm. Web side of Handitek (developing assistive technology for people with cognitive disabilities based on their cognitive abilities). Last visit: January 2006.
4. http://centros6.pntic.mec.es/cpee.alborada/. Web side of Alborada. Last visit: January 2006.
5. IEC 61000-6-3: Electromagnetic compatibility (EMC) – Part 6: Generic standards – Section 3: Emission standard for residential, commercial and light-industrial environments.
6. IEC 61000-6-1:2005. Electromagnetic compatibility (EMC) – Part 6–1: Generic standards – Immunity for residential, commercial and light-industrial environments.
7. IEC 6959-1:2001. Information technology equipment – Safety – Part 1: General require-ments.
8. ISO/IEC 9126-1:2001. Software engineering – Product quality – Part 1: Quality model.
9. ISO/IEC 14598-2:2000(E). Software engineering – Product evaluation – Part 2: Planning and management.
10. ISO/IEC 14598-1:1999(E). Software engineering – Product evaluation – Part 1: General overview.

Developing a TriAccess Reading Environment

Chien-Chuan Cko[1], Ming-Chung Chen[2], Li-Yuan Chen[2], Jun-Nan Guo[1],
Ji-Fu Liu[1], and An-Jun Luo[1]

[1] National Chiayi University, Department of Computer Science and Information Engineering
[2] National Chiayi University, Department of Special Education
300 University Road, Chiayi, Taiwan
{cko, mtchen, lihyuanchen, s0914732, s0914713, s0914742}
@mail.ncyu.edu.tw

Abstract. TriAccess is a web-based integrated reading supporting system. It provides physical, sensory, and cognitive supports for individual students based on their specific limitation or preference. The main idea is to provide multiple means of representation for learners to use various ways of acquiring information and knowledge. In addition, TriAccess also provides curriculum material developers and instructors convenient interface. Curriculum developers could upload text and cognitive supports to a simple web page. Then the webpage will generate automatically the material with needed supports. Instructors could set up the appropriate reading supports and manage the subjects for their responsible students on the web.

1 Introduction

Reading is one of the essential activities for learners to access curriculum. Students learn new knowledge, events or concept through reading the textbooks, reference books and related websites. However, reading is not easy, especially for students with reading difficulties. As a mater of fact, reading is a very complex process. Readers need certain skills and ability to be able to read effectively. Normally, a reader needs his/her hands to hold the book and turn to the target page while reading. This task requires reader's physical capabilities to use both hands. Then he/she needs to "read" the text with eye. This process requires the reader with normal sight and visual perception. Finally, the reader need to be able to recognize and understand the words and comprehends the whole paragraphs, and then the whole articles. The reader needs to illustrate his/her cognitive capabilities and reading strategies. This shows effective reading simultaneously demands reader's competences on physical manipulation, sensory perception and cognitive processing.

In fact, it is difficult or even impossible for many students to read effectively. Students could not read or read with many difficulties, because of their limitations or disabilities, such as visual impairment, dyslexia, tremble, visual perception difficulties, or palsy.

Since reading is very important but difficult for people with disabilities, researchers from various professional fields have engaged in related studies to find out effective

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 839 – 846, 2006.

solutions to help these learners to overcome their reading difficulties. Therefore, there are many assistive technology devices and strategies available now. These devices and strategies were use to help learners to comprehend text by mean of bypassing people's disabilities or augmenting their residual capabilities [3]. Computer is one of the best applications in this field. It allows variation of the appearance and representation of text, and the flexibility of the cognitive supports [4,6]. Computerized reading environment has been on the rise recently. The present study reviewed related literatures and found practical solutions as the followings.

1.1 Physical Access

People with limited physical abilities face reading difficulties at the very beginning step. First of all, people need to maintain their body in a proper position before they could manipulate items appropriately. Then, to interact with computer, they may need to be equipped with control supporting devices, e.g. arm and wrist support system. For people with severe limitations of motor control, they need alternative access solutions, e.g. using mouse equipped with switch, joystick, or infrared mouse to navigate electronic text. When users equipped with adaptive computer devices interact with computer, they usually need some adjustments of interface simultaneously. For example, adjustments such likes larger icon and navigate bar, longer interval time between double clicks, and larger space between icons.

1.2 Sensory Access

It is useful for those people with low vision to magnifying the front size by means of magnifying program, and contrasting the fonts from the background. As for learners who are blind, it is essential to provide refreshable Braille displays or screen reader which could transfer text on demand.

1.3 Cognitive Access

Cognitive access is the main process of reading. However, people with specific learning difficulties, could experience numerous and various obstacles when reading. For example, people with learning disabilities, attention disorder, or mental retardation have restrictions or deficiencies in attention, word identification, and comprehension. The empirical studies have proved some useful strategies in computer program, such as adjusting the space between characters, words, and paragraphs and changing the foreground and the background colors; marking up the keywords or important ideas in the articles; reading the selected text aloud; and providing electronic dictionary, picture symbols assistant, and concept map [4].

The above solutions are effective for assisting readers with diverse capabilities, disabilities, or readers with preferences to understand the content of the article. However, there is no integrated system which could provide physical, sensory and cognitive access in one program. For example, the SeeWord provides various options related to appearance of the text and speech synthesis [4]. The WYNN provides electronic dictionary and highlight tool. The AWB provide dictionary for picture symbols, speech synthesis, and adjustable menu bar [2]. These programs in the market provide only one solution. Another problem for these programs is that some kinds of cognitive supports, such as concept maps and precis, could not be created by adaptive devices automatically. In fact, these cognitive supports could be developed by curriculum professionals instead.

Reader need to access material and access comprehension in reading activity. Recent assisting devices, however, are not available to reach comprehension access. Therefore, it is obvious that supportable reading environment which could provide flexible access methods is in need.

Universal design, originated from architecture, has been a popular idea in device, interface or curriculum design recently [4,5,6]. Inducing from universal design, Universal design learning environment (UDL) which has been advocating by Center for Applied Special Technology (CAST) provide us an important perspective to study the reading problem[1]. There are three major principles of UDL. Firstly, the curriculum should provide in "multi-representations". Then, learners could engage in learning activities with "multi-engagements". Finally, learners could express what they learn with "multi-expressions". "Multiple" means flexibility. That meant designers should embed flexible options in the curriculum material. Then these options could be activated depended on individual needs or requirements.

UDL emphasize that learner should reach the same objects in flexible ways. It is real important for learner who receive challenging curriculum in inclusive education environment. Therefore, this study aimed to develop a supportive reading environment, named as TriAccess reading system, which could provide physical access, sensory access, and cognitive access for readers.

2 TriAccess Reading System

TriAccess reading system is a web-based approach of reading system. The reason for using a web-based approach is not only because it has been an important approach recently, but also because it is convenient for readers, instructors, and curriculum material developers. Through Internet, readers with different reading preferences or needs from different locations to access the same learning material on a specific website; instructors could manage their own responsible students who may come from different schools; curriculum material developers could develop the materials and related cognitive supports in advance and upload these components, then maintain and revise them on the website. Therefore, authors developed the TriAccess in web-based approach.

Authors adopted ASP.NET and JAVA Script as the major tool to develop the user interface and management system. Windows Server 2000 was used to serve as workstation. SQL was adopted to store the curriculum material, user data and parameters related to the settings of learner's reading environment.

In order to integrate speech synthesis systems into TriAccess, authors embedded Microsoft Agent, ActiveX, and Character MP3 (a popular speech synthesized program in Taiwan) as options for learners to select. TriAccess could read aloud whatever the text marked by reader without operating extra speech synthesized program. Besides, TriAccess adopted Stream Media technique to display the video supports more effectively. Furthermore, this system was developed under the rules of Web-Accessibility.

Usability of interface evaluation was conducted when the prototype had been created. Four graduate students who were special education teachers in mainstream school involved in the evaluation. They were asked to operate the interface of instructor and learner to evaluate the ease of interaction and to provide the suggestions of improvement. Later authors modified the interfaces based on the result of their evaluation.

2.1 Major Functions of Accessibility Issue

Authors reviewed the related studies and generated 16 possible solutions of assisting reading comprehension. Then 17 professionals, from curriculum design, special education, and instruction fields, involved in a Delphi survey to rate the importance of 16 adaptations for assisting people to participate reading activity and comprehend what they read. Fourteen items were regarded as important or very important. These solutions were embedded into the TriAccess. These solutions could be grouped into three kinds of accessibility issues based on aforementioned major tasks of reading activity. Here is the brief introduction of the three major functions of TriAccess.

Physical Access. The major features of physical access are flexible menu bar, including adjustable size and space of icons, location of menu bar. The possible target users are those with poor eye-hand co-ordinations.

Sensory Access. The significant features of sensory access are flexible options of appearance of text. It provides adjustable size and color of word, color of background, and space of words and rows. The potential target users are those with low visions and poor visual perception.

Cognitive Access. Providing adaptive cognitive supports is the most critical features of the TriAccess. The supports contain alternated representation to text, multi-representations for explaining the key concepts, precis and background knowledge of the article, and concept map. These cognitive supports could provide the poor readers a suitable assistance when they read.

The aforementioned accessibility improvements will be demonstrated in the description of learner's interface in the next section later.

3 Framework

Based on the multi-users management mechanism and the principles of one article with various adjustable components and displaying assistances meeting learner's needs, TriAccess is designed to include four kinds of user-interface environments in the computer program. The four environments are designed for system manager, material developers, instructors, and learners. The framework of TriAccess is shown in Figure one. The major responsible activities of each user are described below:

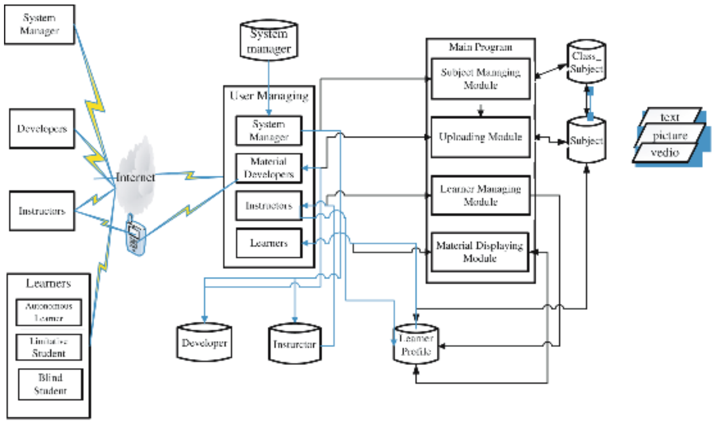


Fig. 1. TriAccess system architecture

System Manager. System manager regulates the profiles of material editors and teachers. System manager create the profile for material editors to let them upload the material for their responsible subjects in advance. When some subjects are available, manager announces the instruction application to the teachers and creates the instructor profiles for the approved applicants that will allow the applicants to use assigned specific subjects.

Material Developers. Although curriculum material developing is a teamwork, TriAccess system does not think that the team members, e.g. instruction, curriculum, psychology, and educational technology, need to developed the integrated web-based reading material by themselves. Instead, TriAccess allows an assigned developer uploads the required cognitive supports separately on a particular web page. Then, the system organizes these supports automatically and stores them in the database.

Instructors. The major responsibility of the authorized instructors is to set up individualized appropriate reading environment for their responsible students. Instructors need to create their responsible students' profiles at first. Then they assign particular subjects for their students. By disabling the unneeded supports, or modify the parameters of the text attribution, instructors adjust the accessibility items for each student individually based on student's abilities and preference.

Learners. Students access particular subjects with adapted reading environment based on their profiles. Based on the consideration of their age, capability of cognition, and experience of computer using, the instructors assign a status for each learn. There are three types of learner status: autonomous, limitative, and blind. Autonomous students are authorized to adjust the accessibility items by themselves. On the contrary, the students with limitative status could only read with the assigned environment. The blind students could use a few hotkeys to access the text. In addition to reading the article aloud, TriAccess could read the explanation of keywords for the blind students.

After learners login in the system and select a specific lesson, they get the individualized reading environment. Examples of screen for learners without blindness shows in Figure two. Menu bar provides learner to select the specific support from the options which is available to him/ her. Then user moves the cursor to a keyword which has been underlined. The related support would display on description area. In voice representation assistance mode, user could mark a specific character, word, sentence or paragraph and click the mouth image on the top to activate text-to-speech program. Computer would speak out the text that is selected. Students with blindness could read the article and the text explanation of keywords in voice by operate some hotkeys.

3.1 Features of TriAccess

Integrating Physical, Sensory, and Cognitive Supports. The activity of reading to learn consists of complex tasks. TriAccess try to integrate physical, sensory, and cognitive supports into a system to provide reading environment as flexible as possible. With in this flexible environment, the problems of providing extra adaptive computer devices will be reduced. For example, poor readers could read with the embedded synthesized program without operating extra speech synthesis system. Besides, TriAccess allows learner not only to “access material” but to “access learning” because of proving suitable cognitive supports. For example, blind learner not only could read the text but the explanation of each keyword by speech.

Flexible Reading Environment with Individualized Supports. As Rose and Meyer indicated, there is no best solution for teachers to design adapted curriculum material for students with various abilities. Instead, providing flexible material may be a practicable solution [6]. TriAccess provides students with various special needs a flexible reading environment to access the same learning material as their classmate without disabilities. Within this environment, learners gain their owned individualized supports which were set up depend on the considerations of learner’s limitations and preferences. For example, for two students who were both struggling with reading text of social studies in a mainstream school, their appropriated assistance may be different because of the different problems. Voice assisting may be useful for one student with poor decoding capability; explanation of keyword may be needed for the other student who lacks background knowledge. Learners could access the same article with different but appropriate supports.

Creating Web-Based Curriculum Material Automatically. TriAccess provides a simple interface for curriculum developers without complex manual procedures. What they need to do is to upload the article and related cognitive supports. TriAccess organizes and creates an article with cognitive supports automatically. In addition, sensory and physical supports options could be generated by system automatically too.

Providing Instructors More Convenient Material Management System. An instructor, especially for special education coordinator or inclusive education supporting teacher, may be responsible for many students with different limitations in different classes or even different schools. It is difficult or hard for them to accommodate material lesson by lesson, student by student, even though the students use the same curriculum. TriAccess allow them to accommodate the supports for each student on the web sites. The supports will be provided automatically to every lesson of all subjects which need accommodation.

Fig. 2. Examples of Learner's interface. The article of these examples is identical but displayed in different conditions and has shown in four frames. The upper left frame reveals picture support of keywords; the upper right one demonstrates video support of keywords; the text of lower left one displays in larger row space and font size; the menu bar and cognitive support displaying area are disabled, and Micro Soft Agent-Merlin is activating.

Passing the Certifications of Web Accessibility. TriAccess is a web-based reading program. It has also passed the priority one requirement by using accessibility evaluation program developed by Government in Taiwan. That means TriAccess could reach minimal accessible requirements. Learners, such as people who are blind, will have no problem to access it.

4 Discussion

This work developed a flexible multi-support reading environment based on the solutions confirmed by related studies. TriAccess provides each learner appropriate physical, sensory, and cognitive supports. Furthermore, TriAccess provides curriculum developers and practical instructors a simple but convenient interface to manage their responsible data. It is an integrated system with multi-supports.

Although each single support was generated from related studies, one question needs further study. That is: is it still useful when embedding with other supports? In the future, we need more clinical experiments to explore the effectiveness of the each support for students with different limitation. Besides, TriAccess provide a platform for curriculum developers instead of providing material development. In Taiwan, only a few material providers (or companies) provide textbook for students in k-12 based on the Standard of National Curriculum. Authors need to cooperate with some of these providers to build content based on their paper-based version in the next step.

Acknowledgment

The authors would like to thank the National Science Council of the Republic of China for financially supporting this research under Contract No. NSC 94-2614-H-415-001-F20.

References

1. CAST (2002): Universal design for access and for learning. Retrieved Nov 3, 2004 from `http://www.cast.org/udl/UDforAccessLearning9.cfm`
2. Chu, C. N., Li, T. Y., Chen, M. C. (2002): The design of an adaptive web browser for young children with reading difficulties. Lecture Notes in Computer Science, 2398, 189-190
3. Lewis, R. B. (1993): Special education technology classroom applications. Pacific Grove, CA: Brooks Cole
4. LoPresti, E. F., Mihailidis, A.., Kirsch, N. (2004): Assistive technology for cognitive rehabilitation: State of the art. Neuropsychological Rehabilitation, 14(1/2), 5-39
5. Orkwis, R. (2003): Universal designed instruction. ERIC/OSEP Digest (E641). ERIC Cleaninghouse on Disabilities and Gifted Education
6. Rose, D. H., Meyer, A.. (2002): Teaching every student in the digital age □Universal design for learning. Alexandria, VA：Association for Supervision and Curriculum Development

A Platform for Creating Adaptive Communicators

María Dolores Paredes Garrido[1], Oscar Pino Morillas[1], María José Rodríguez Fórtiz[1], Manuel González González[2], and Encarnación Rodríguez Parra[3]

[1] GEDES Group. Dpt. Lenguajes y Sistemas Informáticos, Computer Science Engineering Hight School. University of Granada, Spain
{lolapg, mjfortiz}@ugr.es, nosoyblas@hotmail.com
http://www.ugr.es/local/scaut
[2] Santa Teresa Special Education School, Granada, Spain
mayoblan@hotmail.com
[3] Cognitive Ergonomics Group, Dpt. Experimental Psychology Faculty of Psychology. University of Granada, Spain
eparra@cop.es

Abstract. In order to help people with communication problems, we present a platform for creating an augmentative and alternative communicator which runs on a Pocket PC. This communicator is based on an adaptive hypermedia, and may be configured and adapted for each individual person. It has been used with autistic children.

1 Introduction

Augmentative and Alternative Communication systems (AAC) are a growing field of study concerned with providing devices and techniques for increasing the communicative ability of a person with a disability which impedes speech or which makes their communication difficult to understand [1]. There is a wide cross-section of the population who may require an AAC system at any given moment.

We have designed a platform called Sc@ut to create an AAC for children with autism since the communication process is a crucial requirement for their social integration. Autism is considered as a generalized disorder which is characterized both by serious disturbance of certain development areas affecting social interaction and communication skills, and also by stereotypical behaviour, interests, and activities [2]. In addition, 75% of autistic children are also retarded.

Rivière and Martos [3] categorise autism disorders into 11 dimensions. Among themselves, we are most interested in the disorders in children with autism relating to communicative functions, and in particular: Absence of deliberate communication with another person, limited use of words or gestures, serious problems of reciprocity and empathy, problems expressing and understanding language (only 10% use language which is considered normal and receptive language disorders.

AAC [4] systems have been defined as a "structured set of nonverbal codes and their physical support (if applicable) that are learned by the user to allow

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 847–854, 2006.

communicative acts". A traditional classification divides these systems as *AAC systems without assistance* (such as sign language) and *AAC systems with assistance* (i.e. a template or electronic communicator)

AAC tools are fundamental and essential for autistic children's communication and language [5] as an alternative mother tongue [6], and the platform Sc@ut creates an AAC with assistance for children with autism.

Although we are focused in autism, they are many fields where this communicator could be integrated to help to people with another needs of communication, in fact at this moment, we are also working with children who have cerebral paralysis.

This article is organised in the following way: Section 2 compares the desirable features of the Sc@ut communicator with other electronic communicators; Section 3 presents the Sc@ut architecture, distinguishing its two different levels of communicator and meta-communicator, and finally, Section 4 describes our conclusions and future lines of work.

2 Overview of Electronic Communicators

We have studied and compared our proposal with several of the most widely used electronic communicators including *ChatBox, Alphatalker, Deltatalker, Canon communicator, SideKick, Lightwriter SL-20, Cheap Talk 4, TwinTalk* and *VocabPC*.

Basil et al [7] reference to Vanderheiden and Yoder, and they have constructed a list of main features of a communicator. All of the AACs analysed (Sc@ut included) have Vanderheiden and Yoder's standard features. Sc@ut also boasts some of the additional features mentioned, including:

- The use of a Pocket PC to run the communicator
- The design of an application (the meta-communicator) which educators can use to create a communicator for each child.

Mirenda [1] says that a potencial adventage of VOCAs (Vocal Output Communication Aids) is their ability to facilite natural interpersonal interactions and socialization, so Sc@ut can be considered a VOCA system.

3 Sc@ut Architecture

There are two levels to the architecture of the Sc@ut platform [8,9], the lower level is the communicator, the application used by the child, and the higher level is the meta-communicator, the application used by parents and educators to configure the communicator and to adapt it to each child's characteristics. These levels shall be described in the following subsections.

3.1 Communicator

The communicator is a hypermedia system which enables the child to navigate pages with different scenarios and to select components from each page to express what he wants. Each child has their own communicator because of the differences in their abilities to understand language, their capabilities and their skills.

The *scenarios* are constructed by considering the knowledge domain of the child modelled in a conceptual structure. Each scenario comprises elements and represents a real-life action that the child knows and can perform. The elements associated to the action are the scenario *components*.

Each component is located in a cell or square on the hypermedia screen and consists of an image, a text label and sounds. Images of deactivated and unavailable components could also be associated to a component in order to provide information about its state of selection.

A *component's pre-requirements* (such as the mechanical order and/or pedagogical requirements [10]) determine which other components must be previously selected (*and* and *or* logical relationships are established). If a component's pre-requirements do not hold, that component cannot be selected, and this teaches the child to follow an order when selecting.

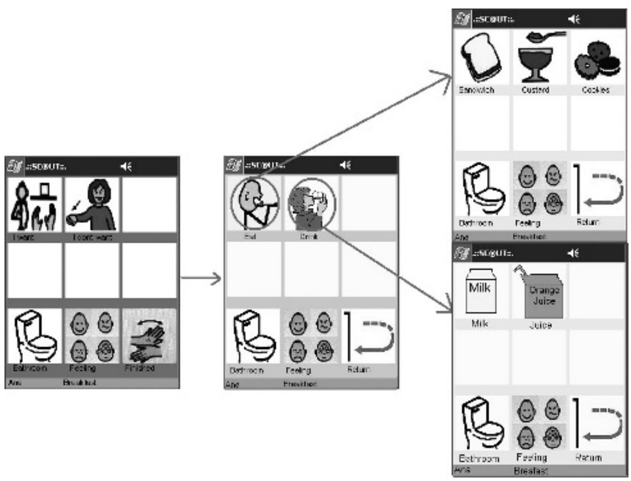


Fig. 1. Sequence of Ana's breakfast

Post-requirements are used to indicate the next components that must be selected or visited. We have seen how certain children are able to construct sentences or can be taught to (even if they do not say them out loud). Post-requirements are used to guide the child in the construction of sentences following the order subject-verb-predicate. The example in Figure 1 shows the hypermedia of a scenario with sentence construction using post-requirements.

3.2 The Meta-communicator

The meta-communicator is the application used by educators to create the communicator for each specific child. There is not a standard configuration to

Fig. 2. Part of Meta-communicator interface

design specific communicators, each communicator is unique because each user is unique, with particular features, and therefore, the communicator must be adapted and adaptable to each child (see Figure 2).

The meta-communicator has been designed to be used by educators who do not have much experience with computers or PDAs. The following steps were taken to create the communicator: (1) Specification of the user profile, (2) Representation of the knowledge domain: scenarios, (3) Creation of the hypermedia model: navigation and (4) Adaptation

3.3 User Profile

Specification of the user profile consists in obtaining each child's particular characteristics that will help to create the communicator, to define navigational aspects, and to adapt the hypermedia model. This information is collected by the child's educators and relatives. Initially, the user profile is presented informally using natural language (see figure 3). At a later stage, it will be used to model the knowledge domain, and it will be structured using behaviour rules which are established on the hypermedia model.

The user profile's characteristics are: child skills, communication habits, guidelines for creating usual scenarios, interaction preferences, educator objectives, interaction shortcuts, and a schedule with the child's main activities.

The way each child navigates the different scenarios also depends on their profile. Different interaction preferences can be chosen in the child's profile such as the number of squares or cells on each page, information mode or interaction type, etc. In this way, two children with the same scenarios and scenario components can both navigate them yet interact differently. Some of these preferences are: *Number of squares* (the number of cells or squares showed on each page); *Information mode* (graphic mode, or text and graphic mode); *Interaction type* (by scrolling or clicking) and if scrolling is chosen as the interaction type, then there are three further characteristics in the user profile: *Previous sound*, *Last minimize* and *Minimum time of answer*; *Repetition of phrase* (the complete sequence of sounds comprising the phrase is repeated once the child has chosen an end-of-phrase component).

In order to implement the user profile, XML templates have been designed since they are easy to create and modify. In addition, if new characteristic data are to be included in the user profile, new labels and values can be added to the XML. The XML values can be used by other modules of the meta-communicator as the creation module for the hypermedia based on this profile.

(1) Personal and communication habits
- *Her name is Ana.*
- *She doesn't talk much but is able to read. She has used templates before to communicate and she is able to create easy phrases with them.*
- *She needs to express her feelings*
- *She often needs to go to the bathroom.*
- *She needs to know when a task has finished.*

(2) Guidelines for creating the breakfast scenario
- *She likes sandwiches a lot.*
- *She also likes chocolate custards but her parents prefer her to eat a sandwich first.*
- *Breakfast doesn't finish until Ana has eaten at least a sandwich.*
- *She likes to drink milk and juice.*

(3) Objectives of her educators
- *She should learn the construction of phrases with subject, verb and predicate.*
- *Her parents want to gradually introduce new foods (e.g. cookies) into her diet.*

Fig. 3. Ana's user profile

3.4 Scenarios and Navigation

The knowledge domain is necessary for modelling the world perceived by the children in a way that is friendly and close to them. As we have seen in the process of constructing the user profile, the educators participate actively to describe the scenarios and activities that the child is able to recognise and understand.

The knowledge domain is represented by means of an ontology. The educators can also define different semantic views of the entire knowledge domain by considering the user profile. The system automatically keeps the integrity of the view. Using the pre- and post-requirements associated to the components for navigating the system, we construct a hierarchy of levels which define how the system is navigated and the syntax of the language for the child. Figures 4a and 4b show two examples of semantic networks for the breakfast scenario for two different children.

If the navigation order of the tree is followed, all the phrases that can be correctly generated are constructed. The arrows define the requirements that the child must satisfy in order to carry out an activity. For example, in Figure 4a, *Ana cannot eat custard until she eats a sandwich or cookies.*

Once the semantic net has been built, the system will be able to automatically generate the interaction templates from it. The templates constitute one of the possible media used to establish communication with the user and are represented in hypermedia format [5]. Pre- and post-requirements allow more complex habits of behaviour to be specified.

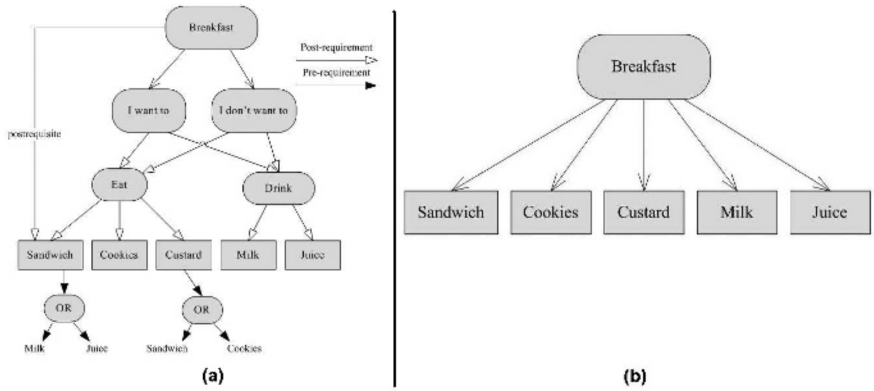


Fig. 4. a) Scenario of Ana's breakfast. b) Breakfast of a child with a low cognitive level.

3.5 Adaptation

The communicator generates *log files* while the child interacts with the communicator. The log file includes information that can be used by the educator to decide to change his profile and his scenarios, and fit it best to the child characteristics. Some information included in the log file is the following: What scenarios and components in the scenarios have been selected; the number of times that each concrete component has been selected, and the number of units initially associated to it; the validity or not of the selection of the component because its precondition has been or not satisfied; the order navigation, path, followed to reach a concrete component from the first template, or from another concrete component.

For the adapting process, *updating rules* based on these observed data are provided [2,6]. If the rules are satisfied, the meta-communicator suggests to the educator what specific changes can be performed. Intervention by the educators is essential for making decisions about the changes since they are the ones who observe the child, who best know their needs and can interpret their responses to the communicator.

In the current version of the communicator, the updating rules have still not been implemented, but the educator can modify certain aspects of the user profile and query the log files information. Adaptive methods and technics will be taken into account during adaptation [13,14,15] and applied in the new version which we are currently working on.

4 Conclusions and Further Lines of Work

In this article, we have presented a communicator for children with autism and also an application for creating and adapting the communicator: the meta-communicator. By dividing the meta-communicator process, concerns can be separated: cognitive, interaction, design and adaptation aspects are differentiated thereby avoiding coupling and enabling the communicator to be adapted for each child. This system has a great variety of adaptive possibilities for user navigation with the communicator, in addition to its educational and pedagogical properties. Sc@ut has been tested with different children with autism and brain paralysis and is currently being evaluated in various schools and so far the results have been satisfactory. We are hoping to be able to provide relevant data relating to this study shortly.

One of the lines of work on which we are currently working is the implementation of a method for updating the rules in order to help in the adaptation process, and the use of integrity rules to decide which changes can be performed at any given moment. Previous research by our group into evolution and adaptation [12] provides the basis for this work.

Another application which will provide educators and parents with necessary backup and support is an online Internet service which they can access in order to download installation files, images, sounds, and tutorials for the communicator. This application must maintain each child's confidentiality and will be accessible during the current year.

One ideal goal might be for there to be a graphic tool to enable the conceptual schema to be drawn, showing relations to each component: images, labels and sounds; as the tool used in previous works of our research group [12,16].

Other future task is to enable adaptations to be made to the current scenario while the child is using the communicator, thus avoiding tantrums. For this, a Wi-Fi connection with the PDA will be used.

Finally, we are researching how to provide the communicator with information about the child's ubiquity in order to automatically load the scenarios, and perhaps some information about the components [17].

Acknowledgments

This research is supported by Sc@ut project (Consejería de Educación y Ciencia de la Junta de Andalucía, Government of Spain) and ADACO project TIN2004-08000-C03-02 (CICYT, Government of Spain).

References

1. Mirenda, Pat.: Toward Functional Augmentative and Alternative Communication for Students With Autism: Manual Signs, Graphic Symbols, and Voice Output Communication Aids. American Speech-Languaje-Hearing Association. **34** (July 2003) 203–216
2. Dautenhahn, K.: Design Issues on Interactive Environments for Children with Autism. Humans & Automation. Seminar Series. (2002)
3. Rivière, Á. and Martos, J.: El Tratamiento del Autismo: Nuevas Perspectivas. IMSERSO. (2001)
4. Schlosser, R.W. and Braun, U.: Efficacy of AAC Interventions: Methodological Issues in Evaluating Behaviour Change, Generalization and Effects. ACC Augmentative and Alternative Communication. **10** (1994) 207–233

5. Tortosa Nicolás F. and Gómez Villa M.: Tecnología de ayuda en personas con trastornos de la comunicación. (2003)
6. Tetzchner, S. v. and Martinsen, H.: Språk og funksjonshemning. Innfring i tegnopplring og bruk av kommunikasjonshjelpemidler. Oslo, Gyldendal. (1991)
7. Basil, C., Soro-Camats, E. and Rosell, C.: Sistemas de Signos y Ayudas Técnicas para la Comunicación Aumentativa y la Escritura. Masson. (1998) 70–72
8. Rodríguez, M.J. Paderewski, P. Rodríguez, M..L. Gea,. M.: An approach for evolving an adaptive and unanticipated system: a communicator for children with autism. FUSE 2004. Workshop on Foundations of Unanticiped Software Evolution, (2004) 40–47
9. García-Cabrera, L., Rodríguez-Fortiz, M.J., Parets-Llorca, J.: Evolving Hypermedia Systems: a Layered Software Architecture. Journal of Software Maintenance and Evolution: Research and Practice. Wiley. **14** (2002) 389–406
10. Hübscher, R.: What's in a Prerequisite. Int'l Conference on Advanced Learning Technology (ICALT 2001). Madison, WI, USA. (2001)
11. Hurtado, M.V., Medina, N., García-Cabrera, L. and Rodríguez, M.L.: An Augmentative Communication System Based on Adaptative Evolutionary Hypermedia System. Lecture Notes in Computer Science **3643** (2005) 223–228
12. Medina-Medina, N. Molina, F, Rodríguez, M.J., García-Cabrera, L.: An Architecture for the Development, Evolution and Adaptation of Hypermedia Systems. Proceedings of the International Conference on Software Engineering Research and Practice. SERP'03. Las Vegas, Nevada, USA, (2003) 112–118
13. Medina-Medina, N.: Un Modelo de Adaptación Integral y Evolutivo para Sistemas Hipermedia. El Sistema de Aprendizaje de SEM-HP. Chapter 3: Adaptive Hypermedia Systems. (2004) 34–36
14. Brusilovsky, P.: Methods and Techniques of Adaptive Hypermedia. User modeling and User-Adapted Interaction, 6. (1996) 87–129. (Reprinted in Adaptive Hypertext and Hypermedia, Kluwer Academic Publishers. (1998) 1–43)
15. De Bra, P.: Adaptive Hypermedia on the Web: Methods, Technology and Applications. Proceedings of the AACE WebNet 98 Conference, Orlando, Florida. (1998) 220–225
16. Medina-Medina, N., Molina-Ortiz, F., García-Cabrera, L.: Diversity of structures and adaptive methods on an evolutionary hypermedia system. IEE Proceedings Software. **152**, Issue 3. (2005) 119–126
17. Gea-Megías, M. Medina-Medina, N., Rodríguez-Almendros, M, Rodríguez-Fórtiz, M.J.: SC@UT: Platform for Communication in Ubiquitous and Adaptative Environments Applied for Children With Autism. Lecture Notes in Computer Science, **3196** (2004) 50–67

Do Text-to-Speech Synthesisers Pronounce Correctly? A Preliminary Study

D.G. Evans, E.A. Draffan, A. James, and P. Blenkhorn

Evans, Draffan and Blenkhorn at School of Informatics
University of Manchester, United Kingdom
{david.g.evans, e.draffan, paul.blenkhorn}@manchester.ac.uk
James at IanSyst Ltd., Cambridge, UK
abi@dyslexic.com

Abstract. This paper evaluates 4 commercial text-to-speech synthesisers used by dyslexic people to listen to and proof read text. Two evaluators listened to 704 common English words and determined whether the words were correctly pronounced or not. Where the evaluators agree on incorrect pronunciation, the proportion of correct pronunciations for the four synthesisers is in the range 98.9% to 99.6% of the 704 words. The evaluators also listened to the same synthesisers speaking phrases in which there were 44 pairs of homographs and determined whether each instance of the homograph was correctly spoken or not. The level of correctness for the four synthesisers ranged from 76.3% to 91.3%.

1 Introduction

Text-to-speech synthesisers are used in a number of assistive technology systems. Although, the best-known application of text-to-speech synthesisers to assistive technologists is probably screen readers for blind people, perhaps the most wide-spread applications are those used to support people with other reading difficulties. This group includes people with learning disabilities, specific learning difficulties (such as dyslexia) and people who are learning a language. There are a number of systems that use speech synthesisers to speak text to a user, often visually highlighting the text on the screen as it is spoken. Examples include[1] Claro Software's ClaroRead Plus [1], Freedom Scientific's Wynn [2], Kurzweil Educational System's Kurzweil 3000 [3] and Texthelp's Read and Write Gold [4].

In this paper, we present our work in the context of a person with dyslexia using such text-to-speech systems to access text. The reason for setting this context, rather than a broader context of disabled people using speech synthesisers, is that the work reported here is the initial part of a larger study to investigate how dyslexic people use speech synthesisers and to determine whether there are issues arising in doing so.

---

[1] This list is not intended to be exhaustive, but merely illustrative of the list of products available.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 855 – 862, 2006.

2 Using Text-to-Speech Synthesis in Assistive Technology

Over the past 20 years speech synthesisers have become ever more widely available for personal computers. Over this time, the quality of the speech has improved greatly. Today's speech synthesisers have a very high degree of naturalness (i.e. they sound much more like a human speaker than their predecessors). However, naturalness is not the only characteristic of a text-to-speech synthesiser; understandability (i.e. the degree to which the text is correctly perceived by the listener) is also of great importance. Naturalness and understandability are, to some extent, orthogonal. A text-to-speech synthesiser may sound exactly like a human speaker, but that does not mean that the text being spoken can be readily understood; conversely, a text-to-speech synthesiser that sounds robotic may have a very high degree of understandability but would never be mistaken for a human speaker. Some blind users of screen readers contend that the more robotic voices are more useful than the more natural voices because they are more consistent and understandable. This observation may be due, in part at least, to familiarity; the listener becoming conditioned to a particular voice. It may also be due to the fact that earlier, more robotic voices may have a more consistent (and less natural) prosody and pronunciation pattern than later systems.

A user with dyslexia will typically use text-to-speech in three situations – for further information see [5]. Firstly, he/she may be trying to understand a body of text, so, for example, he/she will use the system to read a complete document. This is useful to some dyslexic people, whose ability to process text in auditory form, or combined visual and auditory form, is better than their visual processing of text. Secondly, the user may wish to listen to an isolated word. Some users may read text visually and only call upon the text-to-speech system when a problematic word is reached, which requires clarification by listening to it. Isolated words are also spoken when a user is spell checking. In this case the user needs to be informed of the misspelled word and the options for its replacement. Thirdly, text-to-speech systems are used for proof reading. The user checks his/her writing for the correctness of sentence construction, grammar and punctuation. The prosodic variations of modern text-to-speech synthesisers assist in this task.

In order for a user to make full use of the text-to-speech system, the information that is presented must be clear and free from error. In the case of reading a body of text, incorrect pronunciation or other problems (for example, unusual prosodic effects) may distract the user from his/her major task of text comprehension. In the case of isolated word reading, the pronunciation must be correct. This is especially important in spell checking where the user should, if possible[2], be presented with distinctions between the misspelled word and the suggestions for replacement. The important question is, therefore, do text-to-speech synthesisers speak correctly?

---

[2] It may not be possible, for example a misspelling of the word 'fail' may be 'fale', by analogy with the pairings 'hail'/'hale' and 'mail'/male' there may be no distinction between the word and the misspelling. This is also true of homophones that may appear as corrections to the spelling. For example, the misspelling 'wether' will yield both 'whether' and 'weather' as suggested corrections. Note that 'wether' is actually a valid English word, meaning a 'castrated ram', but many spell checkers (including Microsoft Word's) mark it as an error.

3 The Study

Our initial work in this area, and the subject of this paper, was a study to determine the degree to which speech synthesisers pronounced words in an appropriate manner. Two separate trials were undertaken.

a) A list of 704 common English words was listened to in isolation to determine those words that had errors in pronunciation.
b) A set of 44 homographs[3] in the context of complete meaningful sentences.

Four text-to-speech synthesisers were used for this evaluation; they were chosen because they were commonly used by dyslexic people with text-to-speech systems. The synthesisers used were (the names used to refer to the synthesisers in the subsequent text are given in bold):

- AT&T Natural Voices *Audrey*, UK English
- Realspeak *Jane*, UK English
- Microsoft *Mary*, US English
- Plaintalk *Victoria*, US English

For both tests, two evaluators were used, one female and the other male. Both were native English speakers with experience in listening to speech synthesis systems; neither was dyslexic.

3.1 Isolated Word Test

The words were recorded using each of the speech synthesisers into audio files. The evaluators firstly listened to each audio file from start to finish with reference to a paper copy of the list of words. When the evaluator perceived that a word was unusually or incorrectly pronounced, he/she marked his/her list. All the words that were unmarked after this initial run were classified as being correctly pronounced. The evaluator then listened again to all the words that he/she had marked a problematic in the initial run. He/she then made a judgement to classify the word into one of three classes:

- Correct, the word is treated in the same way as those not marked after the initial run
- Incorrect, the word is incorrectly pronounced
- Partially correct, the pronunciation of the word is acceptable, but some characteristic of the pronunciation of the word causes the listener to note that something is unusual. This category is provided as such pronunciations may mislead a user with dyslexia when spell checking or distract the listener when reading a block of text.

An initial summary of the results is shown in Table 1. Table 2 shows the classification of words marked as by both of the evaluators for each of the four text-to-speech synthesisers.

---

[3] Words with same spelling, but different meaning and pronunciation (for example ‘moped’ which may be interpreted as the past tense of the verb ‘mope’ (to sulk) or the noun meaning a motor powered, two-wheeled vehicle with pedals).

Table 1. Summary of Isolated word classification by evaluator and synthesiser

	Female Evaluator			Male Evaluator		
	Correct	Partially Correct	Incorrect	Correct	Partially Correct	Incorrect
Audrey	98.2%	1.7%	0.1%	97.6%	1.1%	1.3%
Jane	99.1%	0.4%	0.4%	98.6%	0.6%	0.9%
Mary	99.1%	0.4%	0.4%	98.6%	0.6%	0.9%
Victoria	97.7%	2.0%	0.3%	96.4%	1.0%	2.6%

The overall level of correctness differs between evaluators, but the overall ranking of text-to-speech synthesisers is consistent Jane and Mary (joint first), Audrey then Victoria.

Table 2. Words marked Partially Correct (p) and Incorrect (i) by each evaluator (F = female, M= male)

Audrey			Jane			Mary			Victoria		
Word	F	M	Word	F	M	Word	F	M	Word	F	M
advertisement	p	p	altruistic	p		advertisement	p	p	advertisement		x
altruistic	p		apologise		p	apologise	x	x	automatics		x
bureaucracy		p	database	x	x	at		p	better	p	
cashier	p	x	dismissal		x	brochure		p	body		x
courier		p	enthusiasm	p		deliberate	p		both	p	x
discrepancy	p		expertise	x	x	despatch	p		brochure		p
exaggerate	p		fiancé	x	x	expertise	p		call	x	x
experienced	p	x	Florida	p		favourite	x	x	caught	p	x
expertise	x		fluctuate		x	from		p	cause		x
general	p	x	satisfactorily		p	hand		p	certificate	p	
glamorous	p	x	subtitles		p	jeopardise	x		certificates	p	
into		x	temperature		p	manufacturer		p	chose		x
itinerary	p		valuable		x	quantity		p	deliberate	p	
large	p	p	young		x	quarter		p	during	p	
quantity		p				recognise	p	p	enthusiasm	p	
secondary		p				revenue		p	example		p
success		x				subtle		p	expertise	p	
term		p				than		x	extension		p
thus	p	x				toward		p	far	p	
town	p	x				with		p	Florida	p	
woman		x				year		x	fluctuate		x
year		p				young		x	for		x
young		x							from		x
									little	p	x
									lose		x
									sceptical	p	p
									she		x
									simple		x
									subtle	x	x
									toward		p
									water		p
									whole		x
									woman		x
									young		x

It is clear from Table 2 that agreement between evaluators on Partially Correct or Incorrect words is relatively rare. The evaluation results can be combined such that errors are counted only when both evaluators indicate an error; where a word is classified as Partially Correct by one evaluator and Incorrect by the other, it is classified as partially correct. The results of this analysis are shown in Table 3.

Table 3. Classification of words for each synthesiser, errors marked only when evaluators agree

	Correct	Partially Correct	Incorrect
Audrey	98.9%	1.1%	0.0%
Jane	99.6%	0.0%	0.4%
Mary	99.4%	0.3%	0.3%
Victoria	99.1%	0.6%	0.3%

As can be seen from this table, the degree of correct pronunciation of common, isolated words by a range of text-to-speech synthesisers is very high.

3.2 Homograph Test

Homographs are words that are written in the same way, but which have different meanings and often different pronunciations – where pronunciation differs they are heteronyms. The degree to which the pronunciation varies is dependent on the heteronym. For example, the difference between the word ‘moped’ a verb (past tense) and as a noun results in different phonemes. In others, it is simply that syllable stress moves, for example the contrast between the noun and verb forms of the word ‘project’ (the stress is on the first syllable for the noun and the second for the verb).

Forty-four homographs were selected by choosing a fairly common and representative sample from the set provided in [6]. The selection of the words was in many ways arbitrary; however, it is argued that this is not important since the aim of the work is to gain some measure as to the degree to which homographs are correctly pronounced rather than to produce results for all.

A sentence or pair of sentences was then constructed that contrasted the different pronunciations of the homograph (for example ‘He *moped*; his *moped* had been stolen’. The sentences were recorded into a single audio file for each of the synthesisers. The procedure for the isolated word matching was followed, with the evaluator marking words that he/she felt were in error, revisiting those considered to be in error and classifying as Correct, Partially Correct or Incorrect. Table 4 shows an initial summary of the results

Again, there is some variation between the evaluators, but the overall ranking of the synthesisers is consistent. There is again variation between the evaluators. Table 5, shows where both evaluators agree.

The combined results of both evaluators are shown in Table 6. In doing so, the results of the final row of Table 5 have been removed. The intention was that ‘supply’ was interpreted in the sense of being supple; however, the sentence can also be interpreted a ‘supply’ in the sense of a source and is thus removed.

Table 4. Initial results for the evaluation of homographs

	Female Evaluator			Male Evaluator		
	Correct	Partially Correct	Incorrect	Correct	Partially Correct	Incorrect
Audrey	87.2%	4.3%	8.5%	81.9%	2.1%	16.0%
Jane	87.2%	5.3%	7.4%	84.0%	2.1%	13.8%
Mary	71.3%	6.4%	22.3%	69.1%	4.3%	26.6%
Victoria	76.6%	10.6%	12.8%	77.7%	2.1%	20.2%

Table 5. Homograph classification where both evaluators agree. Cl. = classification with p = partially correct and x = incorrect. A = Audrey, J = Jane, M= Mary, V=Victoria. ✓ indicates error for that text-to-speech synthesiser.

Word	Context	Cl.	A	J	M	V
overall	He wore a red **overall**	p				✓
entrance	I like to **entrance** people	x			✓	
object	I want to **object**	x			✓	
present	Please **present** me …	x			✓	
record	I would like to **record** this session	x			✓	
refuse	I would like to **refuse** to…	p			✓	
second	We should **second** her to our department	x			✓	✓
		p	✓	✓		
subject	I know that I should not **subject** you to …	x			✓	
approximate	I would like you to **approximate** to it	x			✓	
	The amount is only **approximate**	x				✓
moderate	We need to **moderate** our output …	p			✓	
separate	I think we ought to **separate**	x			✓	
abuse	Don't give that **abuse**	p	✓			
close	I thought the door was going to **close**	x				✓
	That was **close**	x			✓	
diffuse	Particles will not **diffuse** in this atmosphere	p			✓	
house	We are not prepared to **house** him	x	✓			
learned	My **learned** father …	x	✓		✓	✓
bow	I need to **bow** out	x		✓		
	I'll take the red **bow**	x			✓	
invalid	It is **invalid** to call someone an invalid these days	p				✓
	It is invalid to call someone an **invalid** these days	x	✓		✓	
Lead	I need to **lead** you	x				✓
	The compass will be effected by the red **lead**	x		✓	✓	
Live	I like **live** music	x	✓	✓	✓	
moped	His **moped** had been stolen	x				✓
pasty	You are looking rather **pasty**	x		✓	✓	
	It must have been that **pasty** you ate	x	✓			✓
routed	The army was **routed** at the battle …	x		✓	✓	✓
	I then **routed** them via	x	✓			
wound	I **wound** some paper around it	x			✓	
august	I will ask the **august** man to speak	p		✓	✓	
polish	He is so good with boot **polish**	x				✓
wind	This is a **wind** up	x	✓			✓
supply	They move **supply** under pressure	x	✓	✓	✓	✓

It may not be clear from the results in tables 4 to 6 just how good modern speech synthesisers are at resolving homographs when words are given in valid context. To illustrate this, it worth identifying some tests that all synthesisers produced correct results with. These include:

- The wind will *buffet* us on the way to the *buffet* car.
- Don't give me that *abuse*, I do not *abuse* you.
- It was a *moderate* success. We need to *moderate* out output in future.
- I will *read* to you now. Just as I *read* to you yesterday.

The last of these is particularly impressive as the synthesiser determines the tense that is required from other cues in the sentence.

Table 6. Classification of homograph errors for synthesisers, where evaluators agree

	Correct	Partially Correct	Incorrect
Audrey	90.2%	2.2%	7.6%
Jane	91.3%	2.2%	6.5%
Mary	76.2%	4.2%	19.6%
Victoria	85.8%	2.2%	12.0%

4 Discussion

This limited evaluation of speech synthesisers does show that the levels of correctness for single, common words and for homophones in context is very high. This result was somewhat contrary to our preconception that rather more errors would be fond. However, this does not mean that users with dyslexia can successfully use text-to-speech systems to read text and address spelling errors. It simply means that the text is spoken to a very high standard; what is important, however, is how a user perceives and uses this information. The work reported here provides a starting point for further investigation with the knowledge that the text is almost always rendered correctly.

One issue that should be noted is that text-to-speech synthesisers are good at resolving homophones when supplied with context. However, when spell checking the text-to-speech synthesiser will be supplied with a single word (one of the correctly spelt options) with no context. It is impossible to determine how such words should be spoken without their contexts; consider, for example, the word 'read'. The text-to-speech synthesiser must default to one way of saying the word. This may be confusing to the user, who can see the context of the word on the screen and be able to hear the misspelled word in context. Further work is required to examine whether this is a significant issue and, if so, to determine ways in which users can best be supported.

References

1. Claro Software, http://www.clarosoftware.com, Accessed 19 Jan 2006
2. Freedom Scientific, http://www.freedomscientific.com/LSG/products/wynn.asp, Accessed 19 Jan 2006

3. Kurzweil Educational Systems, http://www.kurzweiledu.com/products_k3000win.asp, Accessed 19 Jan 2006
4. Texthelp, http://www.texthelp.com/rwg.asp?q1=products&q2=rwg, Accessed 19 Jan 2006
5. Fiddler, R.: An evaluation of the use of specialist support services by dyslexic students at a higher education institution. Skill Journal, March 2001
6. Higgins, J.: Homographs. http://www.marlodge.supanet.com/wordlist.homograph.html, Accessed 27 Aug 2004

Investigation on Effect of Prosody in Finger Braille

Manabi Miyagi[1], Masafumi Nishida[2], Yasuo Horiuchi[2], and Akira Ichikawa[2]

[1] Graduate School of Science and Technology, Chiba University,
1-33 Yayoi-cho, Inage-ku, Chiba,
263-8522, Japan
miyagi@graduate.chiba-u.jp

[2] Graduate School of Science and Technology, Chiba University,
1-33 Yayoi-cho, Inage-ku, Chiba,
263-8522, Japan
{nishida, hory, ichikawa}@faculty.chiba-u.jp

Abstract. Finger braille is one of the communication methods for the deaf blind, which seems to be the most suited medium for real-time communication by its speed and accuracy in transmitting characters. We hypothesize that the prosody information exists in the time structure of finger braille typing. Prosody is the paralinguistic information that has functions to transmit the sentence structure, prominence, emotions and other form of information in real-time communication. In this research, we performed a cognition experiment on 12 subjects with a simulated output to confirm the effect of prosody in the time structure. As a result, the percentages of questions answered correctly were 79 % for the prosody simulated output and 65 % for the monotonous output. The result shows the possibility that the prosody information of finger braille can be applied to the assistive technologies for deaf-blind people's communication.

1 Introduction

People who have both sight and hearing impairments are known as deaf-blinds; because of their impairments they face many problems in their normal daily life. It is particularly difficult for totally deaf and blind people to acquire vital and sufficient information necessary for daily living, compared with a sighted-hearing person. To obtain information for living, they use tactile sensation instead of auditory and visual sensation. There are several communication means that involves tactile sensation, such as finger braille, manual alphabets, and the print-on-palm method. Deaf-blind people are able to communicate with sighted-hearing people through interpreters. However, there are some problems in such conversations, such as lack of privacy for deaf-blind people and passive-prone information retrieval. How can deaf-blind people have conversations without using interpreters? Assistive technology, such as simultaneous interpreting system and remote meeting system, to support deaf-blind people is strongly required so that they may have conversations without the need for interpreters. We focused on finger braille as the medium to be used in those supporting systems.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 863–869, 2006.

In finger braille, the fingers of the deaf-blind person are regarded as the keys of a brailler (a braille typewriter). A person types the braille code on the fingers of the deaf-blind. Compared with other methods, finger braille is faster in transmitting the codes so it can be adequate for real-time communication. We believe that prosody information included in utterances is the key to real-time communication. In speech, prosody is, for example, information of pitch, rhythm, and pause, which seems to have the function of transmitting information about sentence structure, prominence, meaning, emotions of a speaker and other form of information [1]. Spoken languages seem to employ all types of prosody, which enhance the real-time comprehension of the utterances [2]. We hypothesize that there is similar prosody information in finger braille as well.

In our previous study [3], we analyzed the time structure of finger braille. In this research, we performed a cognition experiment on 12 subjects using with simulated outputs to evaluate the validity of results. In parallel, we attempt to identify prosody information from the typing strength of finger braille. As an example of an application system, the conference system for deaf-blind and sighted-hearing people using a finger braille input/output device is currently proposed [4]. We can apply the result of this research to the system to realize independent conversation of deaf-blind people. Conversation with others is expected to help deaf-blind people to deepen their relationship with others and actively participate in society.

Regarding other works on finger braille, the conditions of optimizing the vibration stimulus on fingers have been investigated for the output system [5]. Moreover, the input system of finger braille has been proposed taking into consideration the acceleration of finger movements [6]. To realize a wearable finger braille interface, the delay time caused by the noise of wireless LAN has been considered in a finger braille interface using wireless connections to a computer [7]. The teaching system that lets a sighted-hearing person to input finger braille has been proposed [8]. A study [9] has revealed that the emotions of the interpreter can be transmitted through changes in the load and duration of typing. However, there is no other study that focuses on prosody that can transmit information regarding sentence structure, prominence, and meaning.

In this paper, we describe finger braille and prosody, and our hypothesis of prosody in the time structure of finger braille in Section 2, experiment of simulated finger braille output in Section 3, the results and discussion in Section 4, and finally our conclusions and future works in Section 5.

2 Method

2.1 Finger Braille and Prosody

In finger braille, the fingers of a deaf-blind person function as a brailler on which an interpreter types braille code, as shown in Fig. 1.

A braille code consists of a combination of six dots. In finger braille, six dots are assigned to the index, middle, and ring fingers of the interpreter's hands. The

Fig. 1. The left side of the picture shows the hands of a deaf-blind and the right side shows the hands of an interpreter. It is common for braille codes to be typed between the tip and the proximal interphalangeal joint of the index, middle and ring fingers.

interpreter presses down the corresponding fingers of his/her hands simultaneously to express a braille code. The Japanese braille code system consists of 46 codes which express kana characters (mora with voiceless consonant), and some special codes. Each country has a six-dot braille code system for their language. It is possible to express those languages with finger braille. Since finger braille is coded similarly to braille, it is easy to apply to digital information devices and equipment.

Interpreters for deaf-blind people transmit much information at one time because they must explain what is going on around them along with other people's utterances. In real-time conversation, the interpreter must transmit the information quickly. One of the features of finger braille is its speed and accuracy in transmitting characters. Between a skilled interpreter and a deaf-blind person, approximately 350 codes can be transmitted per minute. Why can a deaf-blind understand the interpreter even when codes are transmitted one by one so quickly? Another feature of finger braille is its rhythm. Deaf-blind people say that the rhythm of typing differs according to the interpreter. A study [9] has revealed that the emotions of the interpreter can be transmitted through changes in the load and duration of typing. Is information on individuality and emotions included in the prosody of finger braille?

The following requirements for the real-time understanding of utterance were proposed in [2]: a.) the real-time segmentation of meaningful units from continuous input signals is possible; b.) the real-time understanding of relationships between segmented units is possible; c.) the understanding of units at the earliest point of acquisition is possible; d.) the understanding of relationships between units at the earliest point of acquisition is possible; and e.) the identification and understanding of status and emotions of speakers are possible. It is suggested that the key to fulfilling the requirements is prosody. In speech, prosody is, for example, information of pitch, rhythm, and pause, which seems to have the function of transmitting information about sentence structure, prominence, meaning, emotions of a speaker and other forms of information. It is not inherent in speech only. Also in sign language, the prosody phenomenon can be observed in the body movement of a speaker. Any types of spoken language, not only speech but also auditory and visual languages seem to employ prosody, which enhances the real-time comprehension of the utterances. Therefore, there is a possibility of the existence of prosody information in finger braille that is used for understanding and prediction. If the phenomenon and functions of prosody are elucidated, it can be applied to developing a communication system for the deaf-blind.

2.2 Time Structure of Finger Braille

In previous research, we analyzed the time structure of finger braille to clarify the existence of prosody. We instructed skilled interpreters of finger braille to type sentences and recorded the time transition of typing. In our analysis, the duration between the onset of pressure of one typed code and the onset of the next one was defined as the duration of the typed code. The durations of typed codes for one sentence were compared. It was observed that the duration of the last code of each phrase was longer than that of other codes. It was also shown that the duration of the last code of the prominent word and the code just before the prominent word were appreciably longer than the others. These results indicate that the longer duration clarifies the boundary of each phrase or prominent word. From the results, we suggest that time structure of finger braille has a significant role in expressing the structure and meaning of sentences in real-time communication of finger braille.

In other developed finger braille output systems, only one code or one word with a few codes is the output and its recognition rate was investigated to determine the appropriate duration of one code in the output of finger braille. The duration that had the highest recognition rate was set to be the duration of each code. The duration was constant in the output of the systems. The results of the investigations can be useful for beginners of finger braille who recognize each code one by one with accuracy and understand the utterance. However, from the results of our previous research, we observed that the time duration of finger braille typed by interpreters was not constant in communication at a speed almost the same as that in speech. It changes with the structure and meaning of a sentence, which we call prosody phenomenon. To realize a system that supports the real-time communication of deaf-blind people, we believe that the duration of output needs to be changed depending on the prosody information.

2.3 How Can Effect of Prosody Be Measured?

In this research, we conducted an output experiment on a larger number of subjects. We aimed to demonstrate the effectiveness of prosody by adding prosody information to the output of finger braille. However, the level of mastery of finger braille varied markedly among deaf-blind subjects and we concluded that the results could be unreliable. Therefore, we performed a simulation experiment on sighted-hearing people with simulated outputs.

Kana is a Japanese writing system consisting of 46 characters. In the experiment, *kana* characters, which are code level information similar to finger braille, are used as the visual output simulating the volatile one-dimensional characteristic of finger braille. Because finger braille and *kana* characters are both code level media, we conjectured that they are processed by the same language processing even though their modality differs. In our previous research, it was confirmed that the timing structure of finger braille is influenced by the boundary of each phrase, the structure of the sentence, and prominent words. In this experiment, the phenomenon, in which the duration of each code is extended at the end of phrases and sentences was reproduced in the output of *kana* characters.

3 Experiment

First, we performed the preliminary experiment. A *kana* character was displayed at 72 point on the screen of the computer. It was replaced with the next character after the duration of each character. The duration for each character was first set at 500 ms, and then shortened by 100 ms during the experiment. We found that the level of cognitive ability varied from subject to subject. The experiment was performed on the assumption that the subjects could read the characters, so the sufficient output duration for each character reading should be confirmed. However, the difference in the level of understanding between the prosody output and monotonous output at the sufficient output duration was

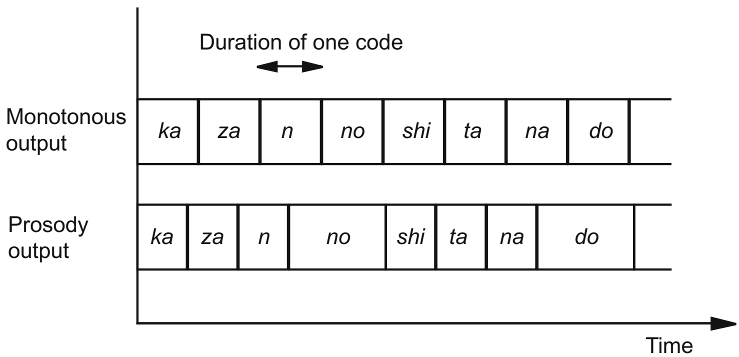


Fig. 2. Example of time structure of outputs. Each box indicates one character. The width of the box indicates the duration of one code. "*kazanno shitanado*" means "under the volcano".

Table 1. Example of output duration. The average duration of whole sentences in the prosody output is the same as that in the monotonous output.

Types of output	Types of code	Time duration (ms)
Determined sufficient time duration		160
Monotonous output		128
Prosody output	End of phrases and sentences	204
	Others	102

not clear. Therefore, we decided to set the output duration at 80 percent of that confirmed above. A shorter duration would place a load on the cognition of subjects so that the difference could be observed in both prosody output and monotonous output.

In the main experiment, six scientific essays (three for output with prosody add-on and the other three for normal monotonous output) with approximately 150 characters were used as the output and four questions were asked for each essay. The subjects were 12 college students. The subjects were instructed to read essays that were displayed only once on the screen, then write the answer for each questions on a sheet of paper provided. The sufficient output duration for each character reading was determined for each subject and we set the output duration at 80 percent of the determined output duration (60 ms to 250 ms). In the output with prosody add-on (hereafter "prosody output"), the duration of characters at the end of phrases and sentences was set to be twice that of other characters. On the other hand, the duration of the normal monotonous output (hereafter "monotonous output") was set to be the same for all characters. (Fig. 2 shows an example of the time structure of both types of output) The average output duration was set to be the same for both the prosody output and monotonous output. In the prosody output, the duration of characters at the end of phrases and sentences was set shorter than that in the monotonous output. Table 1 shows an example of output duration used in the experiment.

4 Result and Discussion

As results of the cognitive experiment, the percentages of correct answers were 79 % for the prosody output and 65 % for the monotonous output. The percentage of correct answers was higher when the prosody output was implemented. It is confirmed that the subjects had a better understanding of the essays given with prosody add-on. The reason for this is considered to be the longer duration at the end of phrases and sentences, which eventually facilitated the segmentation of phrases and sentences for subjects.

In the experiment, it is supposed that the subjects recognized the characters first, and then restructured the characters into phrases or sentences to understand the sentences. However, character recognition in both outputs required much more attention because characters were displayed rapidly from five to ten characters per second. Consequently, the subjects could hardly restructure the characters into phrases or sentences even though they recognized each character. As a result, the percentages of correct answers did not reach 100 %. However, it is suggested that the prosody add-on enhances the understanding even under severe conditions that entail much mental workload.

Even though the duration of standard characters in the prosody output is set shorter than that in the monotonous output, the percentage of correct answers was higher in the prosody output. This indicates that transmitting the structure of sentences enhances understanding even when the time for the recognition of one character is shorter.

5 Conclusion

In this research, we conducted the cognition experiment with a simulated output to confirm the effect of prosody in the time structure. As a result, the percentage of correct answers was higher when implementing the prosody output. The result shows the possibility that the prosody information of finger braille can be applied to the assistive technologies.

In this experiment, characters were output visually, but we will perform the experiment on people with visual impairment who use finger braille, so we can confirm the validity of the prosody of finger braille on many subjects. In parallel, we are currently analyzing the strength in each finger to detect the availability of other possible prosody information in finger braille for real-time communication. In the future, we will implement the addition of the prosody information found from the results of the analysis into the interpreting system to realize real-time conversation between deaf-blind and sighted-hearing people.

References

1. Hirschberg, J.: Communication and Prosody : functional aspects of prosody. Speech Communication, vol. 36. (2002) 31–43.
2. Ichikawa, A.: Dialogue Languages and Persons with Disabilities. IEICE transactions on information and systems, vol. E87-D, no. 6. (2004) 1312–319.
3. Miyagi, M., Fujimori, Y., Horiuchi, Y., Ichikawa, A.: Prosody Rule for Time Structure of Finger Braille. Proceedings of RIAO 2000. RIAO 2000 International Conference, Paris, France (2000) 862–869.
4. Miyagi, M., Nishida, M,. Horiuchi, Y., Ichikawa, A.: Conference System using Finger Braille. Proceedings of GOTHI '05. Guidelines On Tactile and Haptic Interactions, Saskatoon, Canada (2005)
5. Hoshino, T., Kiyosawa, T., Otake,T., Yonezawa, Y.: Basic study of optimum stimulus mode for finger Braille display. The Transactions of the Institute of Electronics, Information and Communication Engineers. vol. J81-A, no. 9. (1998) 1273–1279.
6. Hoshino, T., Otake, T., Yonezawa, Y.: A Study on a finger-braille input system based on acceleration of finger movements. The Transactions of the Institute of Electronics, Information and Communication Engineers. A, vol. J85-A, no. 3. (2002) 380–388.
7. Amemiya, T., Yamashita, J., Hirota, K., Hirose, M.: Virtual Leading Blocks for the Deaf-Blind : A Real-Time Way-Finder by Verbal-Nonverbal Hybrid Interface and High-Density RFID Tag Space. Proceedings of IEEE Virtual Reality 2004 Virtual Reality Conference 2004 (VR 2004), Chicago, IL (2004) 165–172.
8. Matsuda, Y., Sakuma, I., Jimbo, Y., Kobayashi, E., Arafune, T., Isomura, T.: Study on Teaching of the Way to Dot of Finger Braille -Teaching of dotting finger and position of monosyllable-. Journal of Human Interface Society, vol. 7, no. 3 (2005) 379–390.
9. Matsuda, Y., Isomura, T.: Study on Emotional Communication in Skin Contact -Emotional Communication Experiment in Finger Braille-. Journal of Human Interface Society, vol. 5, no. 2 (2003) 163–170.

Internet Chat System for the Deaf-Blind Using Doubled Braille Display – DB4DB

Makoto Kobayashi

Department of Computer Science, Tsukuba University of Technology,
Kasuga 4-12-7, Tsukuba City, Ibaraki, Japan
koba@cs.k.tsukuba-tech.ac.jp
http://www.cs.k.tsukuba-tech.ac.jp/labo/koba/

Abstract. This paper reports on a proposition of a communication system for the blind and deaf-blind person. The system is composed of a terminal which is named "Doubled Braille display terminal for the Deaf-Blind person (DB4DB)." This terminal equipped with two-lined refreshable Braille displays and its software is based on an Internet chat program. The one line of the display is for confirmation of inputted sentence by the user and the other one is for reading all messages from members who take part in the chat space. The proposal system enables the blind and deaf-blind person to recognize message from another member at anytime, even he/she is inputting his/her message.

1 Introduction

In research field of developing a support device for the deaf-blind person, most of researchers have focused on one-to-one communication between the target person and another person, mainly care-giver. For example, the device developed by M. C. Su [1] is a portable communication system that consists of a pair of terminals. One terminal equipped with three Braille cells for the deaf-blind user and another terminal equipped with a Liquid Crystal Display for the other person. Additionally, not only in a research field, but in the consumer market, we can find some other support device using tactile information. Of course I assume that such devices and approaches are truly important for them and basically they need a way to communicate to their care-giver in daily life. However, a viewpoint of my research is slightly different from it.

I think that communication among persons who are in a remote place is as important as communication to the care-giver, in order to keep their QOL. To make it possible, usually, the deaf-blind person utilizes Internet and a personal computer with refreshable Braille display. They can communicate to the remote person through Internet chat program with screen reader which drives the Braille display. It looks like perfect communication tool for them, though, it has one problem.

The problem is that the combination of the Braille display and the screen reader can give the user only one text information from a component where the operating system focused on. In other words, when the other member who participate the chat space says something during the user inputs words, the user has no way of being aware of the message. The deaf-blind user only knows inputting text information. If

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 870 – 873, 2006.

he/she wants to know about the message, he/she has to move the operating focus from the sending text box to the receiving text box. On the other hand, it is difficult to input during he/she reads messages from other participants. The same difficulty was observed when the blind person doing Internet Chat [2].

The problem may cause following uncomfortable situation. The deaf-blind user send inappropriate message when the topic is over. Or other person may hesitate to send a continuous message before getting a reply from the deaf-blind user. Consequently, I can say that this problem strongly prevents smooth communication.

To solve the problem, I tried to make a special terminal for the Internet chat. The terminal and total system is explained in the next chapter.

2 Proposed System Equipped with Doubled Braille Display

2.1 Hardware Description

The idea to solve the problem is very simple. That is to prepare a pair of Braille display arrays for one deaf-blind user. One of the Braille display array is for checking a

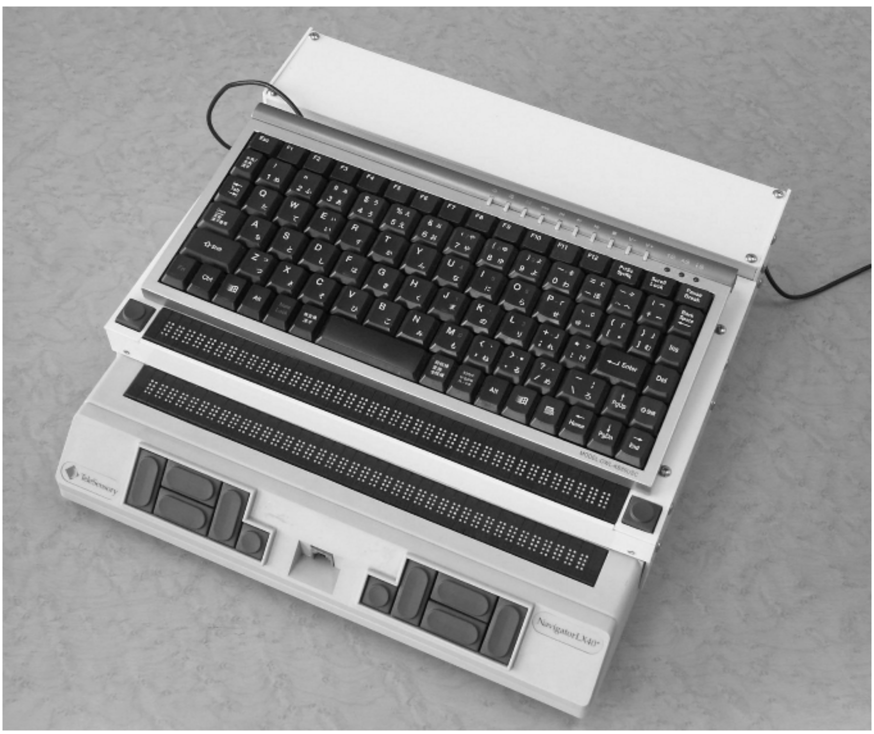


Fig. 1. An overview of Doubled Braille display terminal for the Deaf-Blind person (DB4DB). The Braille array of near side displays all messages from members in a chat space, and another array in back side displays sending message from the user. On the back side of the terminal there are two serial ports to which different control signals are inputted.

sending message and another one is for displaying received messages all the time. The terminal is named “Doubled Braille display terminal for the Deaf-Blind person (DB4DB).” Of course ordinary Braille reading person can not read double line of Braille at a same time, but I think that he/she can recognize the Braille display was changed if the hand touches these displays. Therefore, the user can be aware of the fact of getting a message during input operation.

Identically, the Braille arrays are set on the same surface and it can be realized if they are assigned in face-to-face, though it causes the software control to be complex. Then, I sat an array be simply stacking over to another array shown in Fig.1 as a first step.

2.2 Software Description

To control these Braille display arrays mentioned above, each screen reader program which drives them has to run individually. Preparing two personal computers is one of the solutions, but I selected more convenient and practical way. That is, using "Virtual PC." "Virtual PC" is famous software provided by Microsoft Co. that emulates normal personal computer and we can use several operating systems in one computer hardware.

To install on main operating system and on other operating system in the Virtual PC, two types special chat software was developed. One of them is Chat software for sending and another one is for receiving. The first one only shows editing message to send and the second one only shows receiving message. The user can confirm a message before sending it and can explore the received messages using push buttons

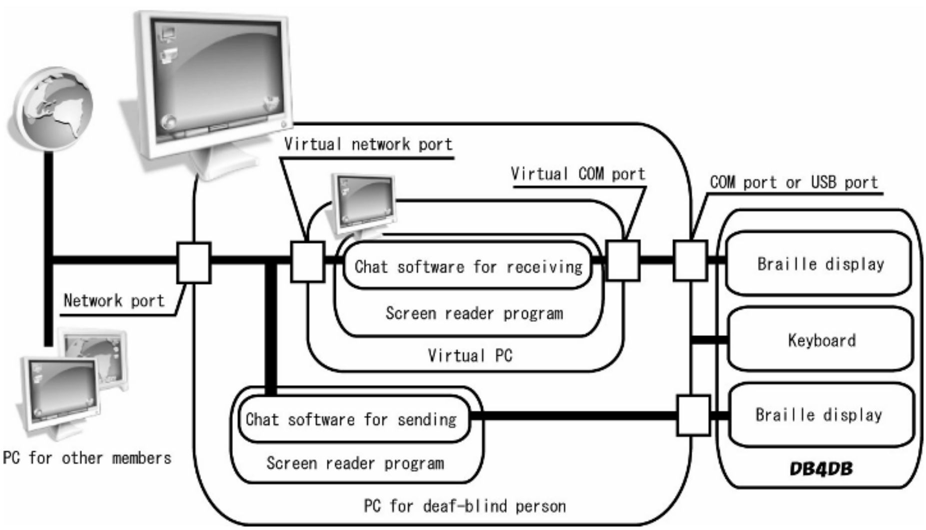


Fig. 2. This figure shows a structure of proposed system. Each Braille display is driven by individual operating system and two operating systems are installed to one computer using Virtual PC. The role of the display (checking a sending message or reading receiving message) is determined by the software.

equipped with the terminal. Fig. 2 shows an overview of the proposed system that explained above. These two programs on different operating system work individually and simultaneously.

3 Summary

In this paper, I proposed a communication system for the deaf-blind user utilizes Internet Chat. A terminal for the deaf-blind user equipped two lined Braille display arrays. It enables the user to recognize received message during input operation is done. The system is not only available for the deaf-blind user, but also for the blind user.

Acknowledgment

This work was sponsored by the Grant-in-Aid for Scientific Research by the Ministry of Education, Culture, Sports, Science (17700466).

References

1. Su, M.C., Chen, C.Y., Su, S.Y., Chou, C.H., Hsiu, H.F., Wang, Y.C.: A Portable Communication Aid for Deaf-Blind People, IEE Computing & Control Engineering Journal, Vol. 12, No.1(2001) 37–43
2. Kobayashi, M., Minagawa, H., Nishioka, T., Miyoshi, S.: Chatting System that Supports Communications between the Visually and the Hearing Impaired, Proceedings of ICCHP2000 Jul.17-21, (2000) 497–502

Using Iconicity to Evaluate Symbol Use

David Gareth Evans[1], Lisa Bowick[3], Marianne Johnson[2], and Paul Blenkhorn[1]

[1] School of Informatics
[2] School of Nursing, Midwifery and Social Work
University of Manchester, United Kingdom
{david.g.evans@, marianne.johnson@postgrad., paul.blenkhorn@}manchester.ac.uk
[3] ENSAD – CNERTA, Dijon, France
Lisa.bowick@educagri.fr

Abstract. This paper investigates the use of iconicity testing to evaluate symbol 'quality' and to examine differences in symbol perception in different ethnic groups. The paper largely replicates an earlier study by Haupt and Alant in which a communication grid of PCS symbols was evaluated with Zulu children. In our study 10 university-educated people with long experience of Western European culture are used to test the symbols. They achieve an overall symbol correctness of 50.3% (compared with Haupt and Alant's 18.9%) and 27.8% symbols are strictly iconic (2.8% for Haupt and Alant) and 55.6 are iconic according to a lenient criterion (11.1% for Haupt and Alant). The concept of distinctiveness as defined by Haupt and Alant is also investigated, as is a method of analyzing symbols based on frequency of selection and correctness when selected. The overall conclusion is that iconicity tests can be usefully employed for assessing symbol quality and determining the difference between ethnic groups.

1 Introduction

There are significant numbers of people who do not speak the official language(s) of the country in which they live. In the UK, for example, there are substantial populations of immigrant communities some of whose members have very limited or not English. This makes many aspects of life difficult, not least in being able to obtain medical treatment and advice. Whilst human interpreters are an important method of militating against the difficulties faced by this group, interpreters are not always available, add considerably to costs and prompt concerns about confidentiality [1]. One proposed approach to supporting communication between healthcare professionals and patients with limited or no English is to use pictographic symbols (of the type used in AAC (Augmentative and Alternative Communication)) [2]. This is being investigated in a pilot project that seeks to evaluate the means by which symbols and computer-based technology can be used to support Somalis with asthma communicate with healthcare professionals. A large proportion of the Somali community have very limited or no English and they may also be illiterate in written Somali. As part of this project it is necessary to evaluate the usability of symbols that

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 874 – 881, 2006.

originate in the AAC community. Ideally symbols would be iconic and transparent, i.e. a Somali user and an English-speaking healthcare professional would use the symbols with no training to communicate with one another. In practice, it seems that using symbols without the support of spoken labels (for the illiterate Somalis) and text labels (for the healthcare providers and the literate Somalis) is impractical [3]. Nevertheless, it seems that symbols have a useful role to play in technology-supported communication; the question is how should one choose the symbols to be used?

There are a number of techniques that are commonly used for the evaluation of symbols. These are: Guessability tests (where participants are asked to guess the meaning of a set of symbols presented independently of one another) [4,5,6]; Translucency tests (where users are asked to rate the match between a symbol and its referent on a discrete rating scale) [7,8,9]; and Transparency tests (where the participant is expected to match one symbol from a small set (generally between 4 and 6) that matches a given referent) [10,11].

For use in medical consultations the symbols must be able to be used with no training of the patients and have a consistent meaning to more than one ethnic group. Whilst translucency testing has been used to compare the perception of symbol by different ethnic groups [9], there are problems with the approach; there is no guarantee that the ethic groups will interpret the scale in the same way. Indeed, our experience is that different groups will use the scales in different ways, hiding their understanding of the meaning. Guessability testing is attractive when no training is required, because it is clear what the user understands by the symbol, and, indeed, such testing is used for safety symbols on medicines where no prior training can be inferred [12]. However, warning symbols on medicines have a well defined context; more generally, context is more variable and Guessability testing can yield to interpretations of symbols, which are, in some ways correct, but which do not match the referent. For example, a traffic light with its green light lit can be interpreted as 'go' (correct) or 'traffic light' (incorrect). Moreover, we argue that in a practical communication system, the selection of one symbol is based not only on the characteristics of that symbol but on the characteristics of the other symbols that the user can select. That is, the selection of one symbol is based not only on the positive features of that symbol, but also on the rejection of all other visible symbols. In this respect, Transparency testing is attractive since it involves the positive selection of one symbol from a set of symbols, the rest of which are rejected. In addition, making a selection of a symbol in response to a referent makes the context clear. In developing a Transparency test, the set of distracter symbols (i.e. those that do not match the referent) is important because their rejection is as important as the selection and they must therefore be realistic in that they are likely to appear as alternatives that a user might select.

An extension to Transparency testing is to use Iconicity testing as proposed by Haupt and Alant [13]. In this test participants are presented with a communication grid with thematically related symbols (Haupt and Alant used a 36-symbol grid related to the task of making a bed). A referent that corresponds to each of the symbols is then read out (in a random order so that each participant's experience is potentially unique) and the participant selects the one symbol from the set that best matches the referent. Thus, for each of the symbols and referents the consistency by which the participants select the symbols can be measured.

Haupt and Alant present results of the Iconicity test with 94 Zulu children aged 10 to 11. The iconicity scores for the symbols are quite low, 2.8% (i.e. 1 of 36 symbols) being iconic on a strict criterion (75% of participants choosing the symbol when its referent was spoken) and 11.1% (i.e. 4 of 36 symbols) being iconic on a lenient criterion (50% of participants choosing the symbol). The overall proportion of correct responses was 18.9%. Haupt and Alant also explore a further classification, which they refer to as 'Distinctiveness', which is the measure of which a symbol corresponds only to its referents and not other referents. We will consider Haupt and Alant's results for this measure when we present our results in Section 0.

Haupt and Alant conclude that the low iconicity scores for their participants may indicate the cultural differences between the Zulu children and the US-derived PCS (Picture Communication Symbols) and grid. In particular, they noted that their participants had problems with symbols that contained arrows.

2 Research Issues

Iconicity testing may prove to be a useful technique for assessing the 'quality' of symbols for users with no training and also be able to indicate the performance of different ethnic groups. The work reported here is a simple pilot study that to a large degree replicates the work conducted by Haupt and Alant with the intention of addressing the following issues:

a. Haupt and Alant report very low levels of iconicity. This may be due to the characteristics of the participants or because the test is intrinsically very difficult. We therefore replicate Haupt and Alant's test with participants whose ethnic background, age and educational level are very different to their participants. If the results differ, it is indicative that the test may provide a sensible means of determining the differences between ethnic groups in their interpretation of symbols. If the results are similar, the test is intrinsically so difficult that the results are meaningless.
b. Haupt and Alant analyse symbol 'quality' by using the concepts of iconicity and distinctiveness. The question is, are these, and possibly other analyses useful measures of the quality of symbols.

3 Method and Results

Ten participants took part; all were educated to at least first degree level from universities in Western Europe, New Zealand or the USA. All used English within a university setting on a regular basis (although for two participants English was not their first language). Six males and four females participated and the average age was 29.8 years (SD = 10.3). No participants had experience of AAC symbols or sign language.

The participants were presented with 36 copies of a test grid made up of a 9 x 4 array of colour PCS. Participants were then presented with 36 referents (with the order randomised between participants) and asked to match each referent to a symbol on the grid. Each time a symbol was matched, the participant used a clean grid to mark his/her selection for the next referent; participants matched some symbols to refer to multiple referents and for all participants some symbols were unused. The symbols and referents are shown in the Appendix together with the number of times they were correctly selected in response to the referent. The symbols used were identical to those used by Haupt and Alant where possible. However, some symbols differ for reasons of availability. Symbols 1, 4, 5 and 18 have quite minor differences. Symbols 20, 24, 27 and 31 have more significant differences. Although, we believe that only symbols 24 and 27 have significant advantages in terms of ease of interpretation. It is believed that the positioning of the symbols on the grid was consistent with that used by Haupt and Alant.

Prior to being presented with the test grid, the participants practised the test using a different grid, which was structured in the same way and used colour PCS. Four referents and hence four copies of the grid were used. The grid consisted largely of nouns and had simple single word referents. All participants correctly matched the referent to the symbol when using the practice grid; different referents were used for each participant.

Overall 50.3% of the symbols were correctly matched to the referents compared with 18.9% in Haupt and Alant's study. Participants were fairly consistent in performance (range 42% correct to 56% correct). The iconicity of a symbol is calculated as the number of correct matches of the symbol to its referent as a proportion of the number of participants. In Table 1, symbols are reported as being strictly iconic where the iconicity of the symbol is ≥ 75% and leniently iconic where the iconicity is ≥ 50%; 10 of 36 (27.8%) symbols match the strict criterion and 20 of 36 (55.6%) the lenient criterion. The strict and lenient criteria are suggested by Haupt and Alant and motivated by [14]. Haupt and Alant's results are given as a comparison in Table 1; Appendix 1 matches symbol numbers to symbols and referent.

Table 1. Iconicity, Strict and Lenient, Haupt and Alant's results taken from [13]

Criteria	Strict (iconicity ≥ 75%)	Lenient (iconicity ≥ 50%)
Our Results Symbol Numbers	3, 10, 11, 14, 16, 17, 22, 24, 27, 32	3, 5, 7, 8, 10, 11, 14, 15, 16, 17, 18, 20, 22, 23, 24, 27, 29, 32, 35, 36
Haupt and Alant's Results Symbol Numbers	11	11, 12, 14, 25

In addition to iconicity, Haupt and Alant also introduce a measure of distinctiveness, i.e. a measure by which the symbol matches to one and only one referent. This is orthogonal to iconicity, and a symbol may be iconic (i.e. it is always selected in response to its referent) but it may also be indistinctive (i.e. it may also be selected in response to other referents). Thus, Haupt and Alant classify their symbols in two dimensions, iconicity and distinctiveness. To present this information, they further relax the iconicity criteria to a more lenient 25% and define a symbol to be distinctive when only one referent accounts for more than 20% of the response for that symbol. We analyse our results in the same way and present them alongside Haupt and Alant's in Table 2. Because our results have much greater levels of iconicity, we show results for an iconicity of 25% (as Haupt and Alant) and also at the 50% level used in Table 1.

Table 2. Iconicity versus Distinctiveness, Haupt and Alant results from [13]. Dist = Distinctive, Indist = Indistinctive.

	Iconicity ≥ 50%		Iconicity ≥ 25%		Haupt and Alant Iconicity ≥ 25%	
	Dist	Indist	Dist	Indist	Dist	Indist
More Iconic	3 5 8 10 11 18 20 22 23 27 29 32	7 14 15 16 17 24 35 36	2 3 5 8 10 11 18 19 20 22 23 27 28 29 32 33	4 6 7 14 15 16 17 24 34 35 36	5 12 16 25 27 35	4 7 13 15 20 26 28 31
Less Iconic	2 12 13 19 25 28 33	1 4 6 9 21 26 30 31 34	12 13 25	1 9 21 26 30 31	11 14	1 2 3 6 8 9 10 17 18 19 21 22 23 24 29 30 34 36

A symbol that is more iconic and distinctive is frequently selected in response to its referent and is not selected consistently for any other symbol. A symbol that is iconic but indistinctive will commonly be used for another referent. For example, Symbol 14 ("Let us make the bed") is iconic according to the strict criteria (in fact it is matched to its referent by all 10 participants), but is indistinct because it is selected another 19 times (for 8 different referents, including 8 times for Referent 21 ("Tuck it in"). Some symbols are more frequently selected in response to the referents for other symbols. For example, Symbol 21 ("Tuck it in") is selected just twice in response to its referent, but is selected 8 times for Referent 12 ("Put it in the Tub") and 7 times for Referent 31 ("You are welcome"), consequently it is indistinct and less iconic. Similarly, Symbol 12 (" Put in the tub") is never selected in response to its referent, but is selected 9 times for Referent 26 ("What a mess") and once each for referents 4 ("You need to change them") and 27 ("It looks like a bomb went off"). It is interesting to note that Symbol 12 would actually be strictly iconic and distinctive for Referent 9.

A further analysis, not performed by Haupt and Alant is to compare the frequency of selection of a symbol with the number of times that the symbol is correctly selected as a proportion of the total number of selections for that symbol. Table 3 shows such an analysis. The columns represent the frequency of selection; high frequency represents more than 12 selections, mid frequency between 8 and 12, and low frequency less than 8. This roughly corresponds to the strict iconicity of 75%. The rows show the degree of correctness; correct shows that more than 75% of the selections of that symbol are in accordance with its referent, partial, between 50% and 75% and incorrect less than 50%. The ideal symbol will have a mid frequency selection (i.e. it is selected as many times as there are participants) and a high degree of correctness. It is not possible for a symbol to have both high frequency of selection and a high degree of correctness. It is interesting to note how the results in Table 2 differ from those in Table 3. For example, symbols 3 ("No"), 8 ("It is nice and clean") and 10 ("It is crooked") are all deemed to be iconic and distinctive in Table 2, yet in Table 3, they fall into different categories. Symbol 3 is only partially correct, as 4 of its 12 selections occur when other referents are spoken. Symbol 8 is correct, but is selected less frequently (6 times correctly and once incorrectly) than the optimum. Symbol 10 is almost ideal, it is selected 10 times, 8 times correctly. This type of analysis can point to a symbol's propensity to be chosen in response to its referent and not to attract selections when other referents are spoken.

Table 3. Frequency versus Correctness of symbols

		Frequency of Selection		
		High	Mid	Low
Correctness	Correct (>75%)	Not Applicable	10 11 20 22 27 29 32	2 8 18 23 25
	Partially Correct (≤ 75%, ≥50%)	17 24	3 5 35	19 28
	Incorrect (<50%)	7 14 15 16 21 33 36	4 6 9 12 34	1 13 26 30 31

The results of our small sample of well-educated westerners show differences to the Zulu children used by Haupt and Alant in their study. The results go some way to upholding Haupt and Alant's contention that the cultural differences between their participants and the intended audience for the symbols pose difficulties. One area of particular difficulty noted by Haupt and Alant was symbols that have arrows. From Table 2, it can be deduced that of the 13 symbols with arrows all are classified as indistinct by Haupt and Alant's participants with 5 being more iconic (4,13,15,28 and 31) and 8 as less iconic (1,3,9,10,19,21,23,29). Our participants note that 7 are distinctive (3,10, 3,19,23,28 and 29, with only Symbol 13 being less iconic). Of the 6 indistinct symbols, 1, 9, 21, and 31 are less iconic and 4 and 15 are more iconic. This may support Haupt and Alant's contention that the arrows are a cultural inhibition. However, it may be that the results are skewed by our participants' generally better performance and that the symbols that involve arrows are trying to communicate more complex concepts. Overall, our participants have a correct response rate of 42% to symbols with arrows compared with an overall figure of 50.3%. By examining Table 1, it can be seen that two symbols (12, "Put it in the tub", and 25, "Puff it up") were iconic (at the lenient level) to Haupt and Alant's participants that were not iconic to ours. This perhaps reflects cultural differences. Certainly none of our participants associated Symbol 12 with a tub and Symbol 25 was selected just once as a correct response to the referent.

4 Conclusions

The work reported here largely replicates Haupt and Alant's study with a very different cultural group and shows that older, well-educated westerners outperform Zulu children. This to some extent supports Haupt and Alant's speculation about the cultural differences interfering with the interpretation of the symbols, although it should also be recognised that our participants were much older. However, the results are useful; before the study we speculated that the task might be so difficult that even our well-educated group might fail to outperform younger subjects. The results do indeed show that comparisons in group performance can be meaningfully drawn.

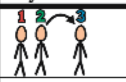
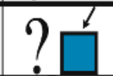
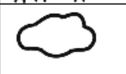
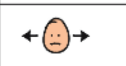
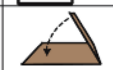
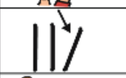
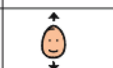
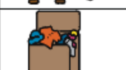
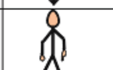
We believe that this type of test is a useful way of determining which symbols are suitable for use with little training. The symbols that are iconic and distinctive or, using the alternative analysis, have mid frequency and high levels of correctness meet the criteria. However, we should acknowledge the fact that the act of choosing a symbol is positive, in the sense that a particular symbol is selected, but also negative, because all other symbols are rejected. Therefore, the performance of a symbol is dependent on the other symbols and referents used. If some symbols are difficult to match to referents, other symbols will be selected in their stead. This means that some symbols that tend to be iconic (because they are selected when presented with the correct referent) may also become indistinctive and their frequency of selection becomes high, not due to defects in that symbol, but to the defects of other symbols. Moreover, it could be argued that symbols that appear to be iconic, distinctive, mid-frequency and correct do so because of the absence of other symbols that might be selected in their place. This factor needs to be taken into consideration so that symbol testing does, to a large degree, match the intended use of symbols in the communication.

Further work will be undertaken with iconicity testing to investigate its utility for determining the 'quality' of symbols in comparison with the other methods noted in Section 1, and to do so with reference to differing ethnic groups.

References

1. Gerrish, K., Ruby, C., Sobowale, A., Birks, E.: Bridging the language barrier: the use of interpreters in primary care nursing. Health and Social Care in the Community, Vol. 12, No. 5, (2004) 407-413.
2. Johnson, M.: What can we learn from drawing parallels between people who use AAC and people whose first language is not English?. Communication Matters, Vol. 18, No. 2, (2004) 15-17
3. Johnson, M. J., Evans, D. G., Mohamed, Z.: A pilot study to investigate communication strategies in provider-patient interaction with Somalis refugees. *Proc. HC2006*, (2006)
4. Cairney, S., Sless, D.: Communication effectiveness of symbolic safety signs with different user groups. Applied Ergonomics, Vol. 13 (1982), 91-97
5. Hanson, E. C., Hartzema, A.: Evaluating pictograms as an aid for counselling elderly and low-literate patients. Journal of Pharmaceutical Marketing and Management, Vol. 9, No. 3 (1995), 51-55
6. Yovetich, W. S., Young, T. A: The Effects of Representativeness and Concreteness on the "Guessability" of Blissymbols. Augmentative and Alternative Communication, Vol. 4, (1988) 35-39.
7. Bloomberg, K., Karlan, G. R., Llloyd, L. L.: The comparative translucency of initial lexical items represented in five graphic symbol systems and sets. Journal of Speech and Hearing Research, Vol. 33 (1990), 717-725
8. Fuller, D. R.: Initial study into the effects of translucency and complexity on the learning of Blissymbols by children and adults with normal cognitive abilities. Augmentative and Alternative Communication, Vol. 7, (1997), 30-39
9. Huer, M. B.: Examining perceptions of graphic symbols across cultures: preliminary study of the impact of culture/ethnicity. Augmentative and Alternative Communication, Vol. 16 (2000), 180-185
10. Musselwhite, C. R., Ruscello, D. M.: Transparency of three communication symbol systems. Journal of Speech and Hearing Research, Vol. 27, (1984), 436-443
11. Mizuko, M.: Transparency and ease of learning of symbols represented by Blissymbols, PCS and Picsysms. Augmentative and Alternative Communication, Vol. 3 (1987), 129-136
12. Dowse, R. & Ehlers, M. S.:. The influence of education on the interpretation of pharmaceutical pictograms for communicating medicine instructions. International Journal of Pharmacy Practice, Vol. 11, (2003) 11-18.
13. Haupt, L., Alant, E.: The iconicity of picture communication symbols for rural Zulu children. South African Journal of Communication Disorders, Vol. 49 (2003), 40-49
14. Doherty, J. E, Daniloff, J. K., Lloyd, L. L.: The effect of categorical presentation on AmerInd transparency. Augmentative and Alternative Communication, Vol. 1 (1985), 10-16

Appendix – Number, Symbol, Referent, and Frequency of Selection

No.	Symbol	Referent	Fq.	No.	Symbol	Referent	Fq.
1		What is next?	1	19		Where is it?	3
2		It is nice and soft	4	20		Look at this	7
3		No	8	21		Tuck it in	2
4		You need to change them	3	22		It is dirty	8
5		Whoops	6	23		Fold it back	6
6		We forgot	4	24		Help me please	8
7		What do you think?	6	25		Puff it up	1
8		It is nice and clean	6	26		What a mess	0
9		Let us take it off	2	27		It looks like a bomb went off	9
10		It is crooked	8	28		Let us do it again	3
11		You need to pull	10	29		Yes	7
12		Put it in the tub	0	30		Put it here	0
13		It is finished	1	31		You are welcome	0
14		Let us make the bed	10	32		The pillow case	9
15		Thank you	5	33		Let me	4
16		The blanket	8	34		It looks bad	4
17		Let us put on	10	35		Hold this please	7
18		The sheets	5	36		It looks good	6

No. = symbol number, Symbol = PCS symbol used, Referent = text to match symbol, Fq. = number of participants who selected the symbol in response to the referent

Development of a Voice-Input Voice-Output Communication Aid (VIVOCA) for People with Severe Dysarthria

Mark S. Hawley[1], Pam Enderby[2], Phil Green[3], Stuart Cunningham[4], and Rebecca Palmer[2]

[1] Dept. of Medical Physics and Clinical Engineering, Barnsley Hospital NHS Foundation Trust, Barnsley UK
mark.hawley@bhnft.nhs.uk
[2] Institute of General Practice and Primary Care, University of Sheffield, UK
{p.m.enderby, r.l.palmer}@sheffield.ac.uk
[3] Dept. of Computer Science, University of Sheffield, UK
P.Green@dcs.shef.ac.uk
[4] Dept. of Human Communication Science, University of Sheffield, UK
s.cunningham@sheffield.ac.uk

Abstract. This paper describes an approach to the development of a voice-input voice-output communication aid (VIVOCA) for people with disordered or unintelligible speech, initially concentrating on people with moderate to severe dysarthria. The VIVOCA is intended to recognize and interpret an individual's disordered speech and speak out an equivalent message in clear synthesized speech. User consultation suggests that such a device would be acceptable and would be useful in communication situations where speed and intelligibility are crucial. Speech recognition techniques will build on previously successful development of speech-based home control interfaces, and various methods for speech 'translation' are being evaluated.

1 Introduction

Many people with disordered speech have intact cognitive and language skills but their disordered speech precludes them from interacting in a manner which allows them to exploit their potential for education, employment and recreation. Dysarthria is the most common acquired speech disorder affecting 170 per 100,000 population [1]. Dysarthria is present in approx 33% of all people with traumatic brain injury, 8-20% of persons with cerebral palsy and increases in incidence with the progression of neurological diseases such as Motor Neurone Disease (MND), Parkinson's Disease and Multiple Sclerosis (MS). Moderate to severe dysarthric speech is often unintelligible to unfamiliar communication partners. As a result, people with dysarthria often use voice output communication aids (VOCAs), but these devices have limitations. Despite much research and development, VOCAs are often slow and tiring to use. Because of their reliance on a switch or keyboard and screen for user interfacing, they do not facilitate natural communication with eye contact between

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 882 – 885, 2006.

conversation partners. Indeed, many people with disordered speech prefer to speak, even if it is effortful and fails, as it is quicker and more immediately responsive than any other communication method.

A communication method is required that retains, as far as is possible, the speed, naturalness and responsiveness of speech communication but, crucially, adds to intelligibility. This paper describes an approach to the development of a voice-input voice-output communication aid (VIVOCA). The VIVOCA is intended to recognize and interpret an individual's disordered speech and speak out an equivalent message in clear synthesized speech.

2 User Views and Requirements

We have consulted VOCA users, including potential VIVOCA users for their views on the acceptability and potential utility of a VIVOCA, and the design features required in such a device. These views are to be incorporated into the design specification.

Eight of the twelve users consulted reported that they would be prepared to use a VIVOCA. They perceived potential benefits to be an increase in the speed of communication and reduced keyboard use; *'I think the idea of a communication aid that speaks a message out clearly would be much easier than typing into a machine', '[it would be] quicker to communicate'*. The users conveyed the idea that the VIVOCA might increase the ability to communicate, increasing self-expression and independence. Specific situations that more than one user would like to use a VIVOCA for include meeting new people, talking on the telephone and shopping; all situations where speed and intelligibility are crucial.

3 Recognition of Disordered Speech

In previous work, we have been successful in developing speech-controlled interfaces for people with severe dysarthria [2]. Our approach has been to apply statistical automatic speech recognition (ASR) techniques, based on hidden Markov models (HMMs), to the speech of severely dysarthric speakers. Viewed as an ASR problem, this task appears initially to be relatively straightforward. Firstly, the input vocabulary was restricted to a small number of isolated words. Secondly, word models were trained for each individual speaker. However, there is only a relatively small amount of data to train the recognition models, and this data exhibits more acoustic variability than normal speech [3]. Conventional HMM training and recognition techniques were used (see [4]), but we have extended the usual methodology by:

- Defining confusability measures which allow us to predict, from a small training corpus, which words are most likely to be mis-recognised [3].
- Developing software which allows the user to practise their productions of a set of words with the aim of increasing their consistency [5,6].
- Using the data collected in these practice sessions to re-train the recogniser, thus closing the loop between speaker training and system training.

This technology (known as STARDUST) has been used to produce a voice controlled environmental control system (ECS) and a basic, small vocabulary VOCA that can be used in the home. In field trials in the homes of users, recognition rates were good for people with even the most severe dysarthria with a mean word recognition rate of 95% under test conditions and 87% in everyday usage in uncontrolled noise conditions.

The current technology has its limitations, however, when applied to the challenging problem of developing effective speech-in/speech-out devices. Firstly, the input vocabulary required for spoken communication is, in general, larger than that required for home control tasks. We aim to discover how far the techniques that have been successful with relatively small vocabularies can be extended to larger vocabularies. Secondly, the VIVOCA device is intended to be used in a wide range of environments, which will present differing acoustic conditions. Although STARDUST applications perform well in the home, performance will diminish where environmental noise levels are higher and more variable, such as in an office or restaurant. Although noise robust ASR remains a major research area, there are a range of techniques that might offer a performance advantage in these conditions. For example, both relative spectral filtering [7] and the missing data techniques [8] will be investigated and, where appropriate, applied.

4 Translating Disordered Input Speech into Synthesised Output Speech

Development of a VIVOCA requires the translation of the limited vocabulary of speech inputs available to people with severe dysarthria into synthesised speech output with a vocabulary adequate to meet their communication goals. We require, therefore, a flexible translation method that is capable of dealing with a variable number of inputs. The intention is to develop a technique combining elements of assistive technology-based solutions with direct speech translation. For users with potentially small vocabularies, we may employ a technique similar to that used in the STARDUST ECS; by combining single utterances with simple grammars a larger number of potential input strings can be translated, using a look-up table, into output phrases. At the other end of the scale, for users with large vocabularies, it may be possible to use direct input word to output word 'translation' for large parts of the desired output vocabulary. Most of the user group will, however, lie in between these two extremes and there are a number of possible translation schemes that we will consider, including word to phrase mapping, pseudo-grammatical combinations and other coding methods. Indeed, there are a number of methods already in existence for producing large vocabulary output from limited inputs, for example semantic compaction [9].

Each of these possible translation schemes will be examined for its suitability in providing the solution we need, in consultation with potential users. Computer models of user translation schemes will be generated to allow comparison of their efficiency, taking into account the variation of speech recognition accuracy with vocabulary size, and the perceptual and cognitive load on the user.

5 User Interface and Hardware Considerations

The ultimate intention is to produce a body-worn device with a user interface that does not require the visual feedback of a screen. In this way, the user can maintain eye contact whilst they communicate. Clearly, the lack of a visual interface is likely to introduce problems for some users in navigating through the process of creating speech communication; therefore we will retain the option for a visual interface. The quality of the microphone is crucial in any speech recognition application and this will be an important consideration in this development. Users have indicated that they want technology that does not 'badge' them as disabled. We intend, therefore, to implement the VIVOCA on standard, commercially available hardware where possible.

References

1. Enderby,P and Emerson, L (1995) Does Speech and Language Therapy Work? Singular Publications p.84.
2. Hawley MS, Enderby P, Green P, Brownsell S, Hatzis A, Parker M, Carmichael J, Cunningham S, O'Neill P, Palmer R, (2003) "STARDUST; Speech Training And Recognition for Dysarthric Users of Assistive Technology." In Assistive Technology – Shaping the Future, eds: GM Craddock et al, IOS Press, pp 959-64.
3. Green, P.D., Carmichael, J., Hatzis, A., Enderby, P.M., Hawley, M. and Parker, M.P. (2003), Automatic Speech Recognition with Sparse Training Data for Dysarthric Speakers, *Proc. European Conference on Speech Technology (Eurospeech)*, Geneva, pp1189-1192.
4. Holmes, J.N and Holmes, W (2001), *Speech Synthesis and Recognition*, Taylor and Francis.
5. Hatzis, A., Green, P.D., Carmichael, J., Cunningham, S.P., Palmer, R., Parker, M.P and O'Neill, P. (2003), An Integrated Toolkit Deploying Speech Technology for Computer Based Speech Training with Application to Dysarthric Speakers, *Proc. European Conference on Speech Technology (Eurospeech)*, Geneva, pp.2213-2216
6. Parker M, Cunningham S, Enderby P, Hawley MS, Green P, (2006) Automatic speech recognition and training for severely dysarthric users of assistive technology – the STARDUST project, Clinical Linguistics and Phonetics Apr-May;20(2-3):149-56
7. Hermansky, H. and Morgan, N. (1994) RASTA processing of speech, *IEEE Trans. Speech & Audio Proc*, 2, 587-589.
8. Cooke, M.P., Green, P.D., Josifovski, L. and Vizinho, A, (2001) Robust automatic speech recognition with missing and uncertain acoustic data, *Speech Communication* 34, 267-285.
9. Cross RT, Baker BR, Klotz LV, Badman AL (1998) Semantic compaction in both static and dynamic environments: a new synthesis, *CSUN conference*, Los Angeles, March 1998. (Available at `http://www.dinf.ne.jp/doc/english/Us_Eu/conf/csun_98/csun98_064.htm`).

EMU – A European Multilingual Text Prediction Software

Christian Beck, Gottfried Seisenbacher, Georg Edelmayer, and Wolfgang Zagler

Institute 'integrated study', Vienna University of Technology
Favoritenstrasse 11/029, A-1040 Vienna, Austria
{beck, seisenbacher, edelmayer, zagler}@fortec.tuwien.ac.at
http://www.is.tuwien.ac.at

Abstract. EMU is a program that supports disabled people to write text faster and/or with less physical load. The program was developed based on knowledge of the outcome of the R&D project IST-2000-25420 FASTY. This paper shows how the commercially available program EMU was developed with special attention to the test results with the final prototype coming from the EC project.

1 Introduction

Motor impairments make the use of standard text input devices to the computer difficult and hence slow. But motor impairment often goes together with articulatory deficiencies. Therefore, communication often relies on slow text-input and that leads to communication disorders and negatively influences the quality of life. Whereas experienced typists, for example, will produce some 300 keystrokes per minute, typical mouth-stick users achieve only around 100 keystrokes and input methods using scanning lower this number to around 3 to 10 keystrokes [2].

EMU assists motor, speech, learning and language impaired persons to produce texts faster, with less physical and/or cognitive load and with better spelling. EMU is highly configurable and available for different European languages. It allows easier access to PC-based office systems, to modern forms of IT communication and a faster usage of text to speech synthesizers for voice communication.

2 The Final FASTY Prototype – The Baseline for EMU

The goal of the EC co-funded R&D project IST-2000-25420 FASTY [1] was the development of software and hardware to ease and speed up the task of writing text with a computer.

The final prototype of the FASTY project shows the following user-interface features:

- Different ways of positioning the text prediction window:
  - Free positioning
  - Following the caret
  - Docked to a certain edge of the screen
  - Docked to an application window

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 886 – 893, 2006.

- Fixed / hardcoded global shortcuts for the prediction selection
- Integrated Adjustment Tool
- Fixed sound and speech output systems with very limited configuration options

The final FASTY prototype uses several methods and algorithms to produce predictions:

- Uni- and bigram-based statistical prediction using general and user frequency dictionaries and a trigram-based Part-of-Speech (PoS) tag model
- Grammar-based prediction analyzing predictions in a wide syntactic context and providing syntactic criteria for ranking predictions
- Collocation based prediction for finding word pairs where one word makes the other one „more likely".
- Compound prediction which allows the creation of compounds „on the fly" without the necessity to store all possible compounds in the dictionaries

3 EMU – The Next Generation

The final FASTY prototype was not ready for the market due to several reasons, some of which – among other things – will be discussed below. Furthermore, many promising enhancements, which were found during the user tests, could not be included in the final prototype within the project runtime. Therefore, a new implementation from the ground, based on the knowledge of the project, was developed.

Based on the existing FASTY knowledge decisions had to be taken, which techniques and features should be used for the development of EMU. The following chapters show some aspects of the design process and the current state of the development.

3.1 The General Structure

The general structure of the FASTY prototype has proved to be well suited for the assignment. Therefore the base concept of EMU was taken from the FASTY prototype (see Fig. 1.).

The structure shows two main parts. The Runtime program is the actual text prediction system. It monitors the user input and offers words or phrases fitting to the text currently written. The Adjustment program is a very powerful and versatile tool for setting up the Runtime program, maintaining the dictionaries and tailoring EMU to the needs of the user in general.

3.2 The Runtime Program

The User Interface. The User Interface is the active part of the EMU Runtime program. It receives keystrokes, initiates the creation of predictions and presents a certain number of predictions to the user. After every typed keystroke a new set of predictions is shown. The user has the choice to type another key (if the correct word was not shown) or to select a word from the prediction list. A screenshot of the currently implemented user interface of the EMU system is shown in Figure 2. The UI may currently be displayed in English, German, Italian and Swedish.

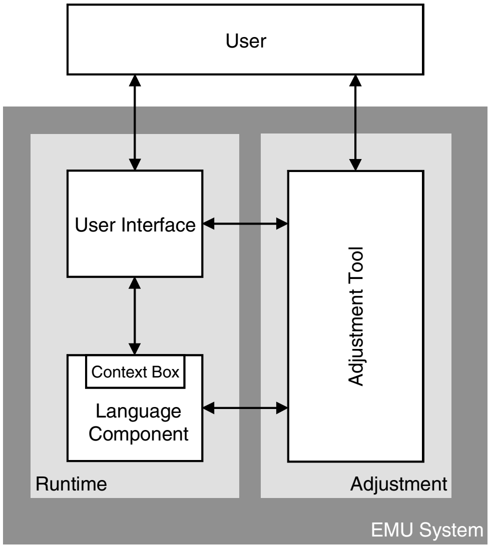


Fig. 1. General Structure of the EMU System

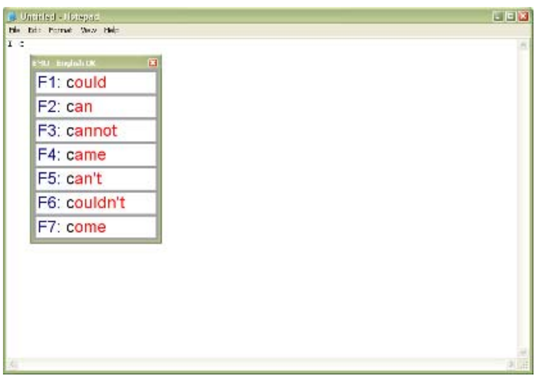


Fig. 2. Screenshot of the EMU Runtime Program: EMU (the small window in the foreground) shows seven predictions for the already written text fragment „I c" within Notepad. The existing character of the currently typed word is shown in black, the possible endings are shown in red; the selection hotkey for each word is shown on the left hand side in blue (all colors may be changed by the user).

The user interface module is more or less platform dependent, especially the methods of catching user keystrokes and putting a selected prediction back to the system by „pressing" the respective keys from the program is strongly connected to a certain Operating System.

The EMU system is currently available for the Windows Operating System only; porting to other systems means implementing the greater part of the user interface anew.

A more detailed view of the user interface reveals, the big differences compared to the final FASTY prototype (see Fig. 3).

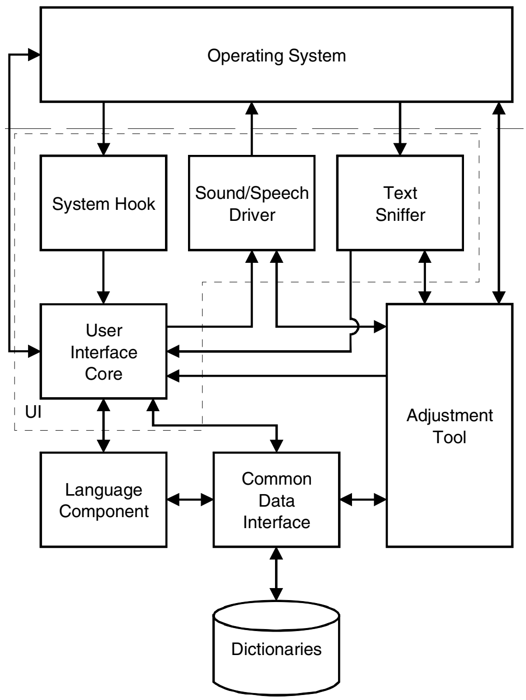


Fig. 3. More detailed structure of the EMU System

The *User Interface Core* is the central part of the UI. All input events are finally received by the Core; text characters are forwarded to the Language Component (see 3.2.2). The Core also starts the generation of predictions and presents a list to the user. If the user selects a prediction the Core initiates the „writing".

Catching user keystrokes is one major task of the *System Hook* module. The module translates different input events to standardized messages, which are then sent to the Core. The second major task of the module is inserting characters into the system. This is totally transparent to other applications; all characters seem to be really typed on the keyboard. Finally, the module also monitors the whole system and reports actions like task-switching, changing input focus or starting/stopping applications. Putting these functionalities into a separate module allows the simple replacement with another module using some different method if necessary.

While output to the screen is done by the UI Core using standard Operating System API calls additional output and feedback by text-to-speech synthesis and different sound events (reading out predictions, indication of the system state through sound output, etc.) is done through a driver system. These *Sound/Speech Drivers* use a standardized interface making it totally transparent to the UI Core.

One strength of the EMU system is the ability to present fitting predictions right after focusing an input control that already contains text. This ability is implemented by so-called *Text Sniffers*. EMU has a default sniffer module for reading text from standard controls. Additional modules may be linked to the program, which allow reading text from applications that do not use standard controls or do not reveal their content in a standard way. The sniffers are also able to find out the read-only state of a control. Again this functionality is offered through a standardized interface. EMU features currently additional sniffers for MS Office and Borland products.

Access to system settings and other data is provided by a *Common Data Interface*. This module is also used by the Language Component and the Adjustment Tool (see 3.3) and offers a centralized interface for data access.

Recapitulating, the structure of the EMU system is designed highly modular with standardized interfaces. In this way enhancements, integration of new output systems or the adaptation to new programs with special needs can be done without changing the whole structure and is therefore highly efficient.

The Language Component. The Language Component is implemented platform independent thus allowing to port it to other platforms without (or at least only minimal) modifications. Furthermore, the module may be switched to another language simply by using another set of dictionaries because the algorithms and methodologies used are language independent [1,6].

The structure of the Language Component (see Fig. 4) is completely tailored to the language independent approach.

The *Context Box* contains the left context of the currently written text. Upon request the Prediction Engine generates a list of predictions fitting to the current context. The list of predictions (also containing additional information like probability or PoS information) is passed back to the caller.

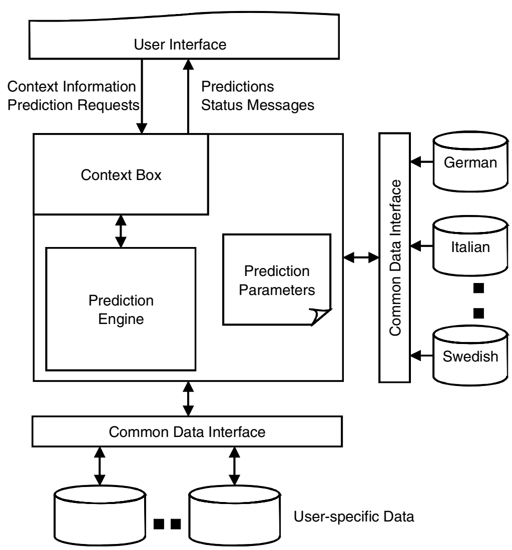


Fig. 4. Structure of the Language Component

The *Prediction Engine* is also able to detect abbreviations in the context and offer the related expansions.

Access to the dictionaries and user-specific data is provided through the *Common Data Interface*. EMU dictionaries are currently available for Dutch, English, French, German, Italian, Spanish, Swedish and Turkish. Most other European languages may be added with limited effort using a set of tools. The original FASTY dictionaries were based on newspaper corpora; for the needs of EMU they were redeveloped to better fit the needs of day-by-day use.

Some prediction techniques of the FASTY prototype where not used in the EMU system, because of the high resource consumption and a quite small benefit. Grammar-based prediction increases the KSR[1] by about 0.2 % points at the cost of a rather high system load [3], the Collocation based prediction increases the KSR by about 0.2% points at the cost of a rather high system load [5]. Future research may change the cost-benefit calculation and the integration of additional prediction modules into the prediction engine can be done quite easily due to the modular structure.

The Language Component of the final FASTY prototype performed within laboratory test environments very well (see Fig. 5), but the subjective feeling of the users was different, because some obviously wrong words made it into the shownprediction list and caused irritation. EMU uses different language data avoiding the very news accentuated predictions, which was the main reason for the wrong words. There are no exact test results up to now, but the user satisfaction (according to user feedback) is much better; the average KSR is some 5% points higher (according to first laboratory tests).

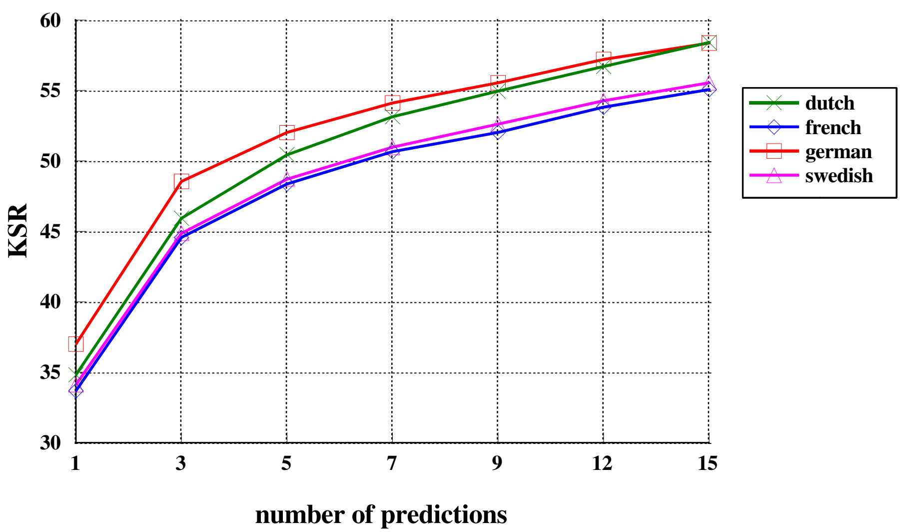


Fig. 5. Result of the tests with the final FASTY prototype

[1] KSR = Keystroke Saving Rate. The KSR is the number of keystrokes saved over the total number of keystrokes.

3.3 Adjustment Tool

The Adjustment Tool (AT) runs independently from the Runtime program and is used to setup and configure the latter easily. Fig. 6 shows a typical screenshot of the Adjustment Tool.

Fig. 6. Screenshot of the Adjustment Tool with open 'Prediction List' page

The Adjustment Tool is totally different to the adjustment options of the final FASTY prototype, because of much negative user feedback during the final user tests within the FASTY project. The EMU Adjustment Tool features very intuitive options and immediate feedback. An example is shown in Fig 6: all font and color options for the prediction window are shown in an example entry on the same page.

The AT uses common interfaces for data access, option storage and data collection. Due to this interface structure changes in EMU subsystems are automatically reflected in the AT without any changes of the AT itself. One example for such an interfaced option page is shown in Fig. 7.

Fig. 7. Screenshot of the Adjustment Tool with open language option page

Acknowledgement and More Information. EMU is based on knowledge and outcomes of the EC co-funded R&D project IST-2000-25420 FASTY (`http://www.fortec.tuwien.ac.at/fasty`).

More information about EMU can be found on the EMU webpage:
`http://www.is.tuwien.ac.at/emu/index_en.html`

References

1. Edited Final Report for IST-2000-25420 FASTY Faster Typing for Disabled Persons, `http://www.fortec.tuwien.ac.at/fasty`
2. Zagler W.L., Seisenbacher G. (2000): "German Language Predictive Typing – Results from a Feasibility Study", Vollmar R, Wagner R (eds.): Proceedings of the 7th International Conference on Computers Helping People with Special Needs (ICCHP 2000), Karlsruhe, Germany, p. 771-779
3. Gustavii E., Pettersson E. (2003) "A Swedish Grammar for Word Prediction", Master's thesis, Uppsala University, Department of Linguistics
4. Baroni M., Matiasek J., Trost H: "Wordform- and class-based prediction of the components of German nominal compounds in an AAC system", Proceedings of the 19th International Conference on Computational Linguistics (COLING 2002), Taipei, Taiwan
5. Baroni M., Matiasek J. (2003): "Exploiting long distance collocational relations in predictive typing", Proceedings of the EACL Workshop on Language Modeling for Text Entry Methods (EACL 2003), Budapest, Hungary, p. 1-8
6. 3rd Edited Annual Report for Publication (IST-2000-25420 FASTY Faster Typing for Disabled Persons), `http://www.fortec.tuwien.ac.at/fasty`

Design and Evaluation of a Versatile Architecture for a Multilingual Word Prediction System

Sira E. Palazuelos-Cagigas[1], José L. Martín-Sánchez[1], Lisset Hierrezuelo Sabatela[1], and Javier Macías Guarasa[2]

[1] Department of Electronics, Polytechnic School, University of Alcalá, Campus universitario s/n, 28805 Alcalá de Henares. Spain
{sira, jlmartin, lisset}@depeca.uah.es
http://www.depeca.uah.es
[2] Speech Technology Group, Department of Electronics Engineering, E.T.S.I. de Telecomunicación, Universidad Politécnica de Madrid
macias@die.upm.es
http://www-gth.die.upm.es/

Abstract. Word prediction is a process that tries to guess the word a user is writing, at the same time he/she is doing it. It is mainly used to decrease the effort needed to write a text in applications devoted to people with disabilities. In this paper, we describe and evaluate the architecture of a multilingual word prediction system. The proposed architecture is modular and flexible, with common interfaces between the modules to allow the use of different prediction algorithms or even the prediction in different languages. The current system consists of a general lexicon for each language, the possibility to create and store personal lexicons, prediction methods based on words and POS (parts of speech) probabilistic grammars (when available). The system has been trained and evaluated for English, Portuguese, Spanish and Swedish. The Spanish version is currently included in a technical aid widely used for people with communication disabilities.

1 Introduction

Generally speaking, a word prediction application tries to find out which is the word a user is typing or is going to type, before he/she writes it completely. The guessed words are shown somehow, so that, if the desired one is included in that list, the user can select and insert it in the text. The time and effort needed to write the text are reduced because the user does not need to type the rest of the word. This is especially useful for slow typists, or people who are not able to use a conventional keyboard. The methods to avoid using a keyboard usually involve the utilization of as many switches as the user is able to handle (usually 1 or 2) and a keyboard emulator, controlled with the switches. Because of the little versatility of the switch-based access, the use of matrices scanning is generally employed. As this is a very slow input method, several acceleration techniques are used to increase the user's typing rate. Word prediction is one of them.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 894 – 901, 2006.

People with other writing problems may also take advantages of the word prediction: dyslexic users or people with frequent spelling errors feel more confident with the help of this aid, and their writing improves in quantity and quality. Children and people learning a second language are also target users of these systems.

2 State of the Art

Several applications based on word prediction have been described in the literature, each one devoted to a particular group of users, with their specific needs (*Profet* [10] *PredWin* [11], *Fazly* [12], *Fasty* [13], *WordQ* [14], etc.). Some of them are speech prostheses to help non-vocal users to communicate: tools including speech synthesis in which user enters text, and it is spoken. Word prediction increases the communication rate of these users, especially if they also have physical problems that make them write slowly. These applications usually include prediction in English, or automatically trained lexicons for other languages. The system explained in this paper supports automatically trained lexicons and prediction methods, but in the case of Spanish, it also has a powerful prediction system with specially designed grammatical information, rules and lexicons that improve the results significantly.

3 General Architecture

In figure 1 the diagram with the general architecture of the system is shown.

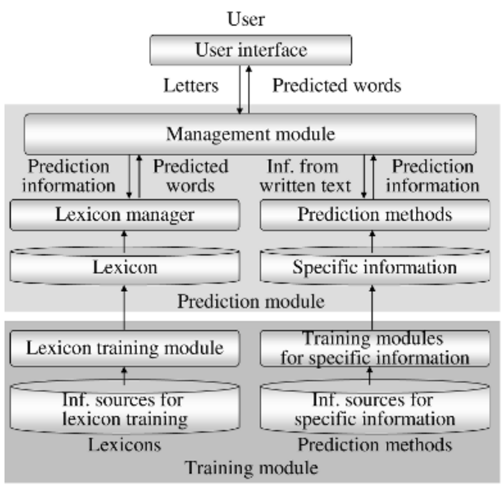


Fig. 1. General architecture of the word prediction system

In the upper part, the modules related to the software application from the user's perspective can be seen, while the lower part is composed of the modules related to the automatic training of all the information sources used in the prediction process.

The prediction methods and the lexicons are independent, making the system design and the portability to other languages easier. Several natural language processing techniques have being used to obtain the best results, but only a subset of them will be used for each language, depending on the available information for it.

The more important modules are described in the following sections.

3.1 Lexicons

The lexicons are the general sources of information of the system. They contain the words and the probabilistic and grammatical information that the prediction methods need to predict each word. There are three kinds of lexicons involved in the prediction process:

- *Main lexicon*, which is a database that contains all the statistical and grammatical information about each word handled in the prediction process: word form, frequency in the training text, and, if available, lemma, part of speech (*POS*), features (gender, etc.), etc. Due to its huge size, hash and other intermediate indexes are used to reduce the computational requirements and to increase the processing speed without loosing information management capability.
- *Subject lexicons*: The system can be trained with selected texts on different topics, in order to generate separate subject lexicons: home, job, music, sports, etc. They contain the word bigrams, and word trigrams for every word in each subject training texts. These lexicons improve the quality of the prediction because the key words of each subject are the most frequent, and need fewer keystrokes to be shown/predicted. Before starting to write, the user should activate the appropriate subject lexicon.
- *Personal lexicon*: It keeps a list of the words and the bigrams written in the present session dynamically, updating its contents at the same time the text is written. It allows the system to adapt immediately to the current subject (maybe there is not a subject lexicon created, or new words appear). This method accelerates the writing mainly because of two factors: the first one is that the user needs to write new words only once (after the first time, they will be predicted as if they were in the main lexicon). The second factor is that word pairs that are typed in a given session are recorded, and each time the first word is written again, the second word is automatically predicted. If one of the possible second words is the intended one, it can be selected directly.

The main problem of the personal lexicon is related to misspelling. Spelling mistakes may be included in the lexicon due to typing errors, very frequent in the use of programs with scanning. These errors are words that shouldn't be included in the dictionary or predicted, so, the inclusion of new words in the lexicon should be carefully controlled. We can think, for example, in a young child learning to write, making spelling mistakes: if the lexicon learns those words, and they are predicted, the child will see them again, and will think they are properly written, getting a counterproductive feedback.

3.2 Word Prediction Methods

The word prediction methods decide the constraints that the words should actually satisfy, from the information provided by the management module and the specific information sources. The methods available in the current version of the system are:

- *Unigram*: based only on the absolute frequency of each word, and predicting the most frequent words of each lexicon.
- *Static bipos and tripos*: based on the probability of the sequences of two or three of parts of speech (*POS*) and the grammatical information included in the lexicons. In the prediction process, the list of words that will be shown to the users are the ones with higher probability according to specific formulae not shown here as it is out of the scope of the paper. If available, basic pos and tripos probability information has been complemented with linguistic features management (that filters the words with not matching features).
- Grammatical word prediction based on a stochastic context free grammar (SCFG), using a specially designed parser described in [3].

3.3 Management Module

It processes the input from the user interface (text written by the user), and manages the information flow between the different prediction methods (coordinating the data each one needs and provides) and the transactions with the lexicons. It obtains the word prediction list that each method provides and combines them to send the most adequate proposals to the user interface.

3.4 Heuristics

In addition to the word prediction methods, a series of auxiliary techniques with a pragmatic base is included in the system architectures to improve the performance of the global system:

- Elimination of the words previously rejected by the user: if a word has been shown to him and, after seeing it, he has not chosen it, the system assumes that it was not the desired one and it will not show it again (until the user starts writing the following word).
- Prediction of the more frequent word suffixes beginning by the last letter.
- Automatic insertion of spaces after punctuation marks and automatic capitalization after a period is written.

4 Description of the Prediction Systems for Each Language

The prediction system architecture is common for all the languages, and only the databases and information sources vary from one to the others. Depending on the sources available at the training stage for each language, there will be (or not) information for some of the prediction methods. For example: if there is no grammatical information in the training test, the prediction methods based on *bipos*, *tripos* and the *parser* will not be used. In table 1 there is a summary of all the information and methods available for each language.

Table 1. Availability of heuristics, lexicons and word prediction methods for each language

Heuristic/Lexicon/ Word Prediction Algorithm	Spanish	English	Portuguese	Swedish
Main lexicon	√	√	√	√
Subject lexicon	√	√	√	√
Personal lexicon	√	√	√	√
Unigram	√	√	√	√
2-grams to 6-grams	√	√	√	
Static bipos and tripos	√	√		
Features management	√			
SCFG	√			
Suffixes prediction	√			
Elimination of rejected words	√	√	√	√
Auto capitalization	√	√	√	√
Spaces after punctuation marks	√	√	√	√

The Spanish prediction system is the most powerful one, because a great amount of information has been created and trained specially for it, and adapted for the prediction process (even the set of grammatical categories was modified to optimize it). The databases for the other languages were automatically generated. 2-grams to 6-grams were not available for Swedish because of lack of a reasonable amount of training text. Static bipos and tripos were trained from tagged text for Spanish and English, but the features management was only available for Spanish. Both the SCFG and the suffixes prediction are only used in the Spanish prediction system, as they were generated manually.

5 Evaluation of the Prediction for the Different Languages

The evaluation of a word prediction system as a writing aid considers its efficiency, measured in two different ways, depending on its application. If it is included in aids for people with physical problems (e.g. to reduce the effort needed to write), we are interested in the saving in the number of keystrokes needed to write the text. If it is included in aids for people with linguistic problems, we should evaluate the "prediction coverage", which can be defined as the number (or percentage) of the words that are correctly predicted (a word is considered as *predicted* if it is proposed by the system before the user finishes typing it). Extended information in the considerations on the automatic evaluation on word prediction systems may be found in [9].

As an example, the results (in percentage of saved keystrokes) of the experiments with different methods, lexicons and languages are summarized in the following tables (the last column shows the relative improvement between experiment *i* and experiment *i-1*). For each language, several experiments have been run: a basic experiment (exp. 1) with the main lexicon and the grammatical information (when available), experiment 2 adding n-grams and experiment 3 adding the personal lexicon to adapt to the new vocabulary in a given topic.

Table 2. Results of the Spanish word prediction system: % of saved keystrokes

Exp.	Configuration	Result	Relative Impr.
1	Static bipos and tripos and features management	42.7%	
2	Exp. 1 plus 2-grams to 6-grams	51.9%	21.5%
3	Exp. 2 plus personal lexicon	53.3%	2.7%

Table 3. Results of the English word prediction system: % of saved keystrokes

Exp.	Configuration	Result	Relative Impr.
1	Static bipos and tripos	28.2%	
2	Exp. 1 plus 2-grams to 6-grams	37.4%	32.6%
3	Exp. 2 plus personal lexicon	47.7%	27.5%

Table 4. Results of the Swedish word prediction system: % of saved keystrokes

Exp.	Configuration	Result	Relative Impr.
1	Unigrams	33.8%	
2	Exp. 1 plus 2-grams to 6-grams	42.7%	26.3%
3	Exp. 2 plus personal lexicon	47.7%	11.7%

Table 5. Results of the Portuguese word prediction system

Exp.	Configuration	Result	Relative Impr.
1	Unigrams	38.2%	
2	Exp. 1 plus 2-grams to 6-grams	42.8%	12.0%
3	Exp. 2 plus personal lexicon	45.0%	5.1%

The differences in the results of experiment 1 (from 28.2% for English to 42.7% in Spanish) are due, not only to the agreement between the test and training texts but also to the amount of information sources available for the particular language. The grammatical information and features management included in the experiment 1 for Spanish have been specially designed to optimize the prediction process, while the grammatical information available for English was the one included in the BNC.

Experiment 2 always improves the results, again depending on the agreement between the test text and the text used to train the n-grams. In case the subject of both texts is the same, the prediction obtained by the n-grams based methods could be very good, leaving a narrow margin to the personal lexicon. This effect can be seen in tables 2 and 5, while in the experiments for English the personal lexicon achieves a relative improvement of over 27%. It is also important to note that, in the case of the best trained languages (Spanish and English), the results for experiment 3 are very similar, showing how powerful the personal lexicon can be (also in Spanish, but the relative improvement due to it is smaller due to the higher performance in the baseline experiment).

The percentage of words predicted before the user types all their letters is usually over 90-95% for all the languages, lexicons and methods.

As it can be seen, the results show that the architecture is able to efficiently handle different languages with similar performance (taking into account the differences between the languages and the training and testing databases, which will be fully described in the final paper).

Additional detailed information about the word prediction techniques, evaluation methods and results can be found in [2].

6 Conclusions

Communication problems have a strong influence in a person's quality of life. The great advances in the telecommunication and computer sciences offer new possibilities that mean big changes in our society, and, particularly, the development of high-tech aids for people with some kind of disability may improve its quality of life.

NLP techniques and, particularly, word prediction, are widely used in technical aids for people with physical disabilities as methods to accelerate the typing speed, increasing the communication rate, and to reduce the effort needed to write a text.

In this paper, the architecture, lexicons and word prediction methods of a prediction system have been described and evaluated for Spanish, English, Swedish and Portuguese. The system architecture is modular and has a high degree of flexibility. It is composed by independent modules with well defined interfaces between them that allow to easily change prediction methods or lexicons.

The evaluation has shown how the architecture is able to efficiently handle different languages with similar results. It is also clear that there are important differences when including additional information sources in the prediction process, when compared with the basic prediction methods. The improvements strongly depend on the previous information (the availability of the grammatical information, features, the amount of words in the main lexicon, etc.). N-grams also produced results varying with the agreement between the test and training texts.

The use of flexible methods, like the personal and subject lexicon, produces the best results for all the languages, due to their capability to adapt to the new vocabulary and frequencies. They compensate the shortages of the fixed lexicons.

The word prediction for Spanish is currently included in PredWin, a widely used text editor for people with severe physical disabilities described in [11].

Acknowledgements

The development of these studies has been funded by the Community of Madrid Biotechnology and Health Science Area, in the project "Global Communication System for people with physical or cognitive disabilities" (Ref. GR/SAL/0814/2004).

References

1. Langley, P. Machine Learning for Adaptive User Interfaces. Proceedings of the 21st German Annual Conference on Artificial Intelligence. Freiburg. Alemania. Springer. http://www. isle. org/langley/adapt. html. (1997). 53-62
2. Allen, J. Natural language Understanding. Benjamin/Cummings Publishing Company Inc 2ª Ed. (1994)
3. Palazuelos S. Aportación a la predicción de palabras en castellano y su integración en sistemas de ayuda a personas con discapacidad física. Tesis Doctoral. (2001)
4. Hierrezuelo L., Palazuelos S. Revisión y propuesta de sistemas de ayuda a la escritura para personas con discapacidad física. Congreso TELEC 04, Santiago de Cuba. (2004)
5. Moulton B., Lesher G., Higginbotham J. A system for automatic abbreviated expansion. http://www.enkidu.net/downloads/papers/MoLeHi99a.pdf. (1999)
6. Demasco, P., McCoy, K. F., Jones, M., Pennington, C., & Snyder, S. Compansion: A technique that applies natural language processing to augmentative communication. Rehabilitation R & D Progress Reports / Journal of Rehabilitation Research and Development, 29. (1991).180-181.
7. Karttunen L. Applications of Finite-State Transducers in Natural-Language Processing. http://www.xrce.xerox.com.
8. Earley, J. An Eficient Context-Free Parsing Algorithm. Comm. of the ACM. Volumen 13. No 2. (1970). 94-102.
9. Palazuelos S. E., Aguilera S., Rodrigo J. L., Godino. J., Martín J. Considerations on the Automatic Evaluation of word prediction systems. Augmentative and Alternative Communication: New Directions in Research and Practice. Whurr Publishers. London. (1999). 92-104
10. Carlberger A., Carlberger J., Magnuson T., Palazuelos S., Hunnicutt S., Aguilera S. Profet, a New Generation of Word Prediction: an Evaluation Study, Workshop on NLP for Communication Aids, ACL/EACL'97. Madrid. España. Proceedings de Workshop on NLP for Communication Aids. ACL/EACL'97. 1-908-873-3898. Ed. "Assotiation for Computational Linguistics. Madrid, España. (1997) 23-28
11. Palazuelos S., Martín J., Arenas J., Godino J. Aguilera S. Communication strategies using PredWin for people with disabilities. Conference and Workshop on Assistive Technology for Vision and Hearing Impaired. Castelvecchio Pascoli, Italia. (2001).
12. Fazly A, Hirst G,. Testing the Efficacy of Part-of-speech Information in Word Completion, in 10th EACL, Proceeding of the Workshop on Language Modeling for Text Entry Methods, Budapest (Hungary), (2003).
13. Gustavii Ebba, Petterson Eva. "A Swedish Grammar for Word Prediction" http://www.stp.ling.uu.se/~matsd/thesis/arch/2003_gustavii_pettersson. pdf.
14. WordQ Writing Aid Software. http://www.academicsupport.net/word-prediction.htm

People with Motor and Mobility Impairment: Human Computer Interaction, Rehabilitation

Introduction to the Special Thematic Session

Michiaki Yasumura[1], Kazuyuki Itoh[2], Akira Ichikawa[3], and Akira Okamoto[4]

[1] Faculty of Environmental Information, Keio University,
5322 Endo, Fujisawa, Kanagawa 252-8520 Japan
yasumura@sfc.keio.ac.jp
http://web.sfc.keio.ac.jp/~yasumura/indexe.shtml
[2] National Rehabilitation Center for Persons with Disabilities,
4-1 Namiki, Tokorozawa, Saitama 359-8555, Japan
ito@reha.go.jp
http://www.rehab.go.jp/english/index.html
[3] Graduate School of Science and technology, Chiba University,
1-33 Yayoicho, Inage-ku, Chiba 263-8522 Japan
ichikawa@faculty.chiba-u.jp
http://www.ailab.tj.chiba-u.jp/en/home.html
[4] Research and Support Center on Higher Education for the Hearing and Visually Impaired,
Tsukuba University of Technology,
4-3-15 Amakubo, Tsukuba, Ibaraki 350-8520 Japan
okamoto@k.tsukuba-tech.ac.jp

Abstract. Here we describe the summary of the STS on “People with Mobility Impairment: Human Comuter Interaction, Rehabilitation” that was proposed by Japanese research group of the Grant-in-Aid Scientific Research of the Japanese Ministry of Education, Culture, Sports, Science and Technology. Twelve excellent papers are presented, whose authors are from Europe, North America, and Asia. Among them, 6 papers focus on cursor control by eye movement, and 6 papers describe other new challenges of assistive technologies.

1 Introduction

This STS is proposed by Japanese research group of the Grant-in-Aid Scientific Research of the Japanese Ministry of Education, Culture, Sports, Science and Technology "Basic Research on the Communicative Functions of the Elderly and Persons with Disabilities".

The STS provides a forum for the discussion of major issues related to the topics above in order to promote usability and accessibility of information and communication equipment by the elderly and persons with disabilities.

It is essential that the elderly and persons with disabilities be able to use information and communication equipment without any difficulties in order to access necessary information and to communicate with other persons easily in their daily life.

Recent progress of ICT (Information and Communication Technology) has made equipment highly sophisticated, small and inexpensive. New kinds of easy-to-use

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 902 – 904, 2006.

equipment have also been developed and the advancement of assistive technology is remarkable. As a result, the use of ICT equipment by the elderly and persons with disabilities is increasing. However, the usability and accessibility of ICT equipment is still insufficient for them. To make ICT equipment easy to use for the elderly and persons with disabilities, it is necessary to form the basis of information accessibility by conducting basic research which focuses on how they feel, what they think, how they act and how they behave towards the equipment. Also, data relevant to machines easy for them to operate should be collected and analyzed.

2 Outline of This STS

This STS focuses on “People with Motor & Mobility Impairment”, which is one of the topics of the above research area. Twelve excellent papers are presented, whose authors are from Europe, North America, and Asia. Among them, 6 papers focus on cursor control by eye movement, and 6 papers describe other new challenges of assistive technologies.

Chin (U.S.A) found that their method for pointing using EMG signal from facial muscle can get faster point-and-click time than the old system.

Lotte and Struijk (Denmark) devised a tongue-computer interface that uses an activation unit, with inductors placed in the user’s mouth cavity. Using inductive coupling, rather than a force activated switch, makes activation faster and less fatiguing.

Chen and Wo (Taiwan) designed an alternative Chinese keyboard layout for a single-digit typist and compared its typing speed and keystroke accuracy with the conventional QWERTY key layout and other layouts.

Herrera et al. (U.S.A) investigated the design and implementation of a writing module that uses myoelectric signals to control the griper on a prosthetic hand to help amputees write characters. Writing characters by 7-segment strokes was successful.

Takahashi et al. (Japan) proposed “Assistance type handrail” by which an individual’s motor functions could be fully maximized. One Parkinson subject could stand up with their power-assisted handrail.

Inoue et al. (Japan) developed a driving simulator that consists of a PC and a haptic device (steering wheel and switches) for rehabilitation. The results of their preliminary experiments with brain injured subject showed its effectiveness.

Hori et al. (Japan) made several improvements on a traditional EOG system to develop a real-time EOG communication control system for developmentally disabled individuals with motor paralysis. They examined the usability of their input system with EOG and a virtual keyboard and got acceptable results.

Fejtová et al. (Czech) introduces a device that enables non-contact control of a PC through eye (or head) movement. A tiny camera, which monitors eye movements, is attached to a frame of a pair of glasses in order to make it easy to install and operate.

Daunys and Ramanauskas (Lithuania) developed and analyzed a pupil/eye corner tracking system to find the best mapping from image coordinates to computer screen coordinates. The results of the experiment showed that there are nonlinear distortions between image coordinates and screen coordinates.

Wo and Chen (Taiwan) studied a systematic method to investigate the effect of keyboard adaptation on children with cerebral palsy. In an experiment using their developed keyboard adaptation assessment program, it was illustrated that the typing speed of the participants improved after adaptation.

Itoh (Japan) made a prototype of a light-spot-operated mouse emulator. To control the cursor, two cameras are set in front of the screen and detect light spots on the screen sent from laser pointer mounted on the user's cap or sun visor. The subjects of the GUI operation experiment felt comfortable using this system.

Lin et al. (Taiwan) designed a chorded keyboard which focused on input speed and learning time for people with physical disabilities. The learning curve for the use of the proposed chorded-based input method was very short.

3 Significance of This STS

Research and development of assistive technologies for elderly and persons with disabilities has been widely conducted and several useful results have contributed to the improvement of their QOL and ADL. However, still these results cannot cover the diversity of needs posed by persons with disability, especially the physically impaired. Another problem to be pointed out is the insufficient research effort on users' mental or psychological aspects. Fundamental research on the problem of how they feel, what they think and how they act towards assistive technology should be encouraged.

The research results to be presented in this STS will be a steady advancement in this area and contribute to solve above problems.

Even with disabilities, the life of an individual, whether an older person or a person with disabilities is irreplaceable. To make these precious lives profitable, assistive technologies are necessary. This STS will provide opportunities for researchers to seriously discuss problems they may have, and will yield fruitful outcomes.

Electromyogram-Based Cursor Control System for Users with Motor Disabilities

Craig Chin, Armando Barreto, and Miguel Alonso Jr.

Florida International University
10555 West Flagler Street
Miami, FL 33174, U.S.A.
{cchin006, barretoa, malon05}@fiu.edu

Abstract. An improved hands-free cursor control system suitable for use by individuals with spinal dysfunction or spinal cord injury is introduced. The system uses electromyogram (EMG) signals from facial muscles to produce five distinct cursor actions, namely: left, right, up, down and left-click. The new system is derived from a system previously created by our group. Object selection tests are performed on both systems. We use statistical analysis and Fitts' law analysis of these tests to support our assertion that the new system provides enhanced performance over its predecessor.

1 Introduction

Computer-based systems have become increasingly pervasive in every arena of human activity. Many professions require access to computer-based applications in order for an individual to fulfill job requirements. With the advent of the Internet, the computer has become a portal to a new domain of social interaction, information access, and entertainment. Typically, able-bodied individuals interact with a computer using standard input devices, such as, a mouse, trackball, touchpad, or keyboard. However, there is a considerable segment of the population (e.g., 250,000 – 400,000 individuals in the United States) who are often unable to use such input devices because they live with spinal cord injury or spinal dysfunction [8]. There exists a need to provide such individuals with a more usable means of computer access than these input devices.

With today's GUI-based PC software, most of the human-to-computer interaction is based on selection operations, which consists of two steps:

- Pointing: Positioning the cursor at the desired location of the screen, over the appropriate area or icon.
- Clicking: Executing the Mouse Down/Up function that is interpreted by the computer's operating system as an indicator to complete the selection of the item associated with the icon at the location of the screen cursor.

The cursor control system that we have created facilitates these point-and-click operations through the detection and processing of electromyogram (EMG) signals.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 905 – 912, 2006.

Electromyography is the study of muscle function through monitoring of the electrical signals generated by the muscle [4]. A surface electrode placed on the skin above a superficial muscle will receive electrical signals emanating from several muscle fibers associated with different motor units. The spatio-temporal summation of these electrical signals results in what is called an EMG signal. Therefore, the EMG signal provides an effective means of monitoring muscle activity.

EMG signals have been used previously for cursor control. EMG-based systems have been used in [5,9] and [2,3], with [2,3] focused specifically on the use of EMG signals from cranial muscles. Monitoring the EMG signals of cranial muscles makes this approach suitable for individuals suffering from severe motor disabilities and who are also paralyzed from the neck down.

EMG-based cursor control systems have been shown to perform slowly when compared to a mouse-operated system in object selection tests [2,3]. However, EMG-based systems have the advantage of allowing for small cursor movements suited for high resolution computer displays. This is a quality that is not possessed by other alternative cursor control approaches, such as some based on eye-gaze tracking (EGT) [1,6,7].

The EMG system developed previously by our group [2,3] utilized three electrodes that measured EMG signals from muscles in the head of the user. The EMG signals were classified into cursor actions by performing real-time spectral analysis of these signals. A previous empirical study of the EMG signals from different muscles revealed that they possessed distinguishing frequency characteristics. An example of this is displayed in Fig. 1.

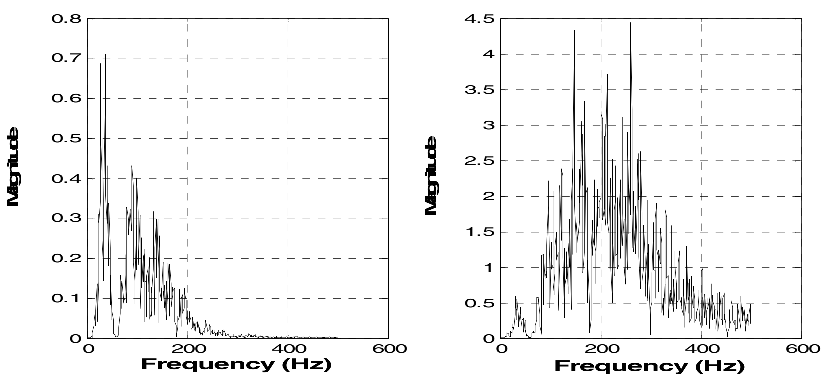


Fig. 1. Spectra observed during a right frontalis contraction (left plot) and a left temporalis contraction (right plot)

After a thorough evaluation of the previous EMG system, it was found that it was occasionally inaccurate in discriminating between the muscle contractions that command up and down cursor movements (eyebrows up and eyebrows down, respectively). To remedy this problem an additional electrode was added to the forehead region, and a new classification algorithm was devised to work with this new input configuration.

Section 2 of this paper details how the new system was implemented and the methodology behind the new classification algorithm. Section 3 also describes the experiment used to obtain object-selection task times, as well as, the data analysis methods used to study these task times. Section 4 provides the results of statistical analysis and Fitts' Law analysis performed on the experimental results. Section 5 presents our conclusions.

2 System Implementation and Signal Processing Methodology

2.1 Placement of Electrodes for the New EMG-Based Cursor Control System

Fig. 2 displays the placement of the Ag/AgCl electrodes on the head of the subject. This figure indicates that electrodes were placed over the right frontalis muscle, the left temporalis muscle, the right temporalis muscle, and the procerus muscle, respectively. An electrode was placed over the right mastoid as a reference. This electrode set up differs from the previous EMG-based system only in the addition of the fourth electrode over the procerus muscle.

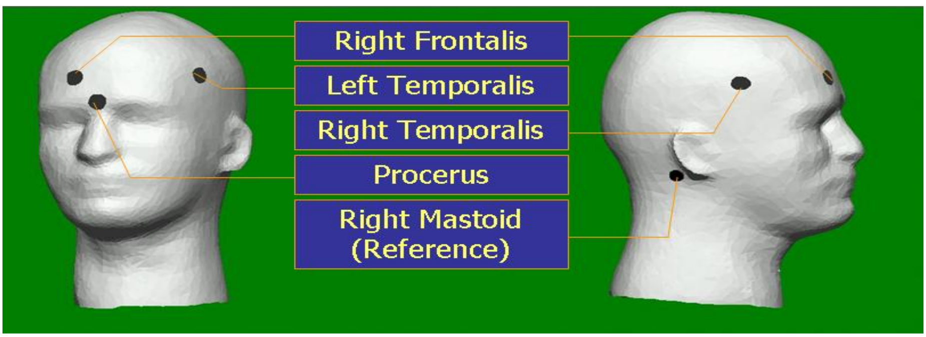


Fig. 2. Electrode placement for the new EMG cursor control system

2.2 Hardware Components of the EMG-Based Cursor Control System

The hardware components of the cursor control system are presented in Fig. 3. Each of the four EMG signals was magnified by a Grass® P5 Series AC preamplifier. Each of these preamplifiers possessed an anti-aliasing filter with a gain of 10,000 V/V and a 60Hz notch-filter. The ADC64TM DSP/AD board (Innovative Integration, Simi Valley, CA) performed analog-to-digital conversion on each signal at a sampling rate of 1.2 kHz, and then applied the classification algorithm to these digitized signals in real-time. The output of the board was a series of TTL-compliant binary voltage sequences consistent with voltage sequences expected from a serial mouse. The Motorola® MC1488C RS-232C driver converted the TTL sequences into RS-232C format and transmitted these sequences into the serial port of the personal computer (PC). The serial mouse driver of this computer communicated with the operating system to produce cursor actions driven by the serial signal created by the DSP board.

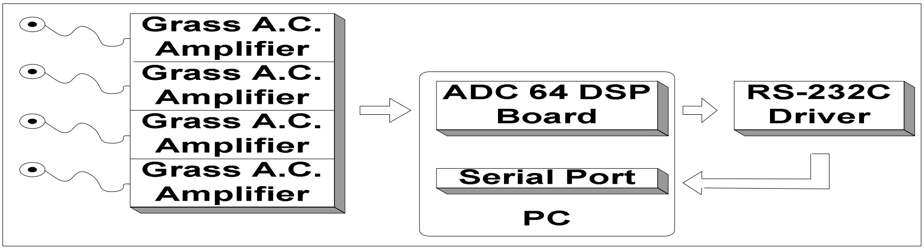


Fig. 3. Block diagram of hardware components of EMG-based cursor control system

2.3 EMG Processing Algorithm for Muscle Contraction Identification

The desired relations between cursor actions, facial movements, and muscle contractions are given in Table 1.

Table 1. Relations between cursor actions, facial movements and muscle contractions

Cursor Action	Facial Movement	Muscle Contraction
Left	Left Jaw Clench	Left Temporalis
Right	Right Jaw Clench	Right Temporalis
Up	Eyebrows Up	Right Frontalis
Down	Eyebrows Down	Procerus
Left-Click	Left & Right Jaw Clench	Left & Right Temporalis

The purpose of the classification algorithm was to determine if a facial muscle contraction had occurred and if so, which specific muscle was the source of this contraction. Given the one-to-one correspondence between muscle contraction and cursor action, the output of an effective muscle contraction classification algorithm can be utilized in a real-time implementation for hands-free cursor control.

Both the classification algorithm of [2,3] and the new classification algorithm made use of the periodogram estimation of the power spectral density (PSD) of the input EMG signals. In both cases, the PSD indicated how the power of an EMG signal was distributed over a frequency range of 0 Hz – 600 Hz. Periodogram PSD estimations were taken every 256 consecutive samples (every 0.213s) from each of the EMG channels.

The two classification algorithms differed in the way each utilized the PSD estimates to classify the EMG data. The algorithm of [2,3] only utilized three electrodes placed on: the left temporalis muscle, the right temporalis muscle, and the right frontalis muscle respectively. This algorithm calculated partial accumulations over the frequency ranges of 0 Hz – 145 Hz and 145 Hz – 600 Hz of the PSDs produced from the three EMG channels in order to distinguish between the frequency characteristics associated with the contraction of different muscles (temporalis versus frontalis). This algorithm also utilized PSD amplitude thresholds to estimate the strength of contraction from each of the three muscles mentioned previously.

Testing of this algorithm revealed that it did not always classify the eyebrows down movement efficiently. So it was proposed that an additional electrode be placed over the procerus muscle, because it is one of the muscles directly involved in the eyebrows down facial movement. This new four-electrode input configuration required a new classification algorithm, the details of which are described in the following paragraphs.

The new classification algorithm made use of Mean Power Frequency (MPF) values to distinguish spectral differences associated with each facial muscle contraction, instead of partial PSD accumulations. The MPF is derived from the PSD values as a weighted average frequency in which each frequency component, f, is weighted by its power, P. The equation for the calculation for the MPF is given by:

$$MPF = \left(\frac{f1 \times P1 + f2 \times P2 + ... + fn \times Pn + ... + fN \times PN}{P1 + P2 + ... + Pn + ... + PN} \right). \qquad n = 1,2, \ldots, N. \qquad (1)$$

The new classification algorithm used a combination of PSD amplitude thresholds, complete PSD sums, and MPF values to correctly classify muscle contractions. For example, to correctly classify a unidirectional muscle contraction all the following criteria must be satisfied:

1. The maximum PSD amplitude must exceed the threshold set for that electrode.
2. The sum of the PSD amplitudes for the given electrode must exceed the PSD sums of the other electrodes
3. The mean power frequency calculated from the PSD must fall into a range consistent with the muscle associated with the electrode (frontalis: 40 Hz – 165 Hz, temporalis: 120 Hz – 295 Hz, procerus: 60 Hz – 195 Hz).

Detection of a click action requires fulfillment of similar requirements in both temporalis muscles simultaneously.

3 Testing and Data Analysis Methods

The experiment used to determine the point-and-click capabilities of the two systems followed a 2 x 4 x 4 x 6 factorial design with two cursor control systems (old and new), 4 different “Start” icon positions (Upper Left, Lower Left, Upper Right, and Lower Right), 4 different “Stop” icon sizes (8.5 x 8.5 mm, 12.5 x 12.5 mm, 17 x 17 mm, 22 x 22 mm), and 6 different able-bodied male subjects. A program was created in Visual Basic to present the point-and-click interface to each subject and to record the movement times required for each task. The program was displayed on a 17” color monitor. For each point-and-click trial, an 8.5 x 8.5 mm “Start” icon was presented in a corner of the screen and a “Stop” icon was presented in the center. Each subject was instructed to use the EMG-based cursor control system to click the “Start” button to begin timing a trial, move the cursor to the “Stop” button, and click on it as quickly as possible. This would record the total task time for the trial. The subject would then click a “Next” icon to display another trial layout with the “Start” button located in another corner of the screen. Each specific trial configuration (“Start” location, “Stop” size, and algorithm) was repeated three times by each subject resulting in a total of 96 trials executed by each subject.

The statistical analysis involved applying a four-way analysis of variance (ANOVA) of this factorial experiment. The results are given in the following section. Fitts’ law analysis applies Fitts’ law to object selection data such as those produced by our experiment. Fitts' law states that there is a linear relationship between the movement time taken for a point-and-click task (MT) and the difficulty of this task (ID). The Fitts' law equation is given by:

$$MT = a + bID \ . \tag{2}$$

where MT is the movement time in seconds, ID is the index of difficulty for the task, in bits. Also, a (seconds) and b (seconds/bit) are the coefficients associated with the linear relationship. ID is given by:

$$ID = \log_2 \left(A/W + 1 \right) . \tag{3}$$

where A represents the distance to the target, and W represents the width of the target, which in our case is the "Stop" button.

Fitts' law essentially states that the narrower and further away a target is, the more difficult the task will be and the more time it will take to be completed.

For Fitts' law analysis, a movement time value is obtained by averaging all the movement times taken for a task of a given ID. Provided that there are tasks with different ID values then we will have a number of (ID, MT) ordered pairs. These ordered pairs are used to produce a linear regression line that represents the performance capabilities of that cursor control system. More specifically, the reciprocal of the slope of the regression line is used as performance measure. The name of this measure is the index of performance (IP) and has units of bits/s. The IP value for a point-and-click system indicates the rate of user information processing for that system.

4 Results

The analysis of variance of the data performed in Minitab produced the following table:

Table 2. Four-way ANOVA table

Source	DF	SS	MS	F	P
Algorithm (A)	1	10290.1	10290.1	84.66	0.000
"Start" Position (P)	3	611.3	203.8	1.68	0.171
A*P	3	221.2	73.7	0.61	0.611
"Stop" Icon Size (I)	3	3073.4	1024.5	8.43	0.000
A*I	3	306.5	102.2	0.84	0.472
P*I	9	833.8	92.6	0.76	0.652
A*P*I	9	603.8	67.1	0.55	0.836
Subject (S)	5	27421.5	5484.3	45.12	0.000
Error	539	65514.2	121.5		
Total	575	108875			

Table 2 shows a significant main effect for algorithm (A) with $p < 0.0005$, and a significant main effect for icon size (I), $p < 0.0005$. Therefore we can reject the null hypothesis of H_{o1}: $A_1 = A_2 = 0$, as well as, the null hypothesis H_{o2}: $I_1 = I_2 = I_3 = I_4 = 0$.

The mean point-and-click task times obtained from the real-time experiment were 22.66 s for the old system and 14.21 s for the new system.

The Fitts' law analysis data are shown in tables 3 and 4.

Table 3. Aggregated point-and-click data for Fitts' law analysis of old system

D(mm)	W(mm)	ID(bits)	MT(s)	IP=ID/MT(bit/s)
180	8.5	4.47	26.87	0.166
180	12.5	3.94	24.24	0.163
180	17	3.53	19.74	0.179
180	22	3.20	19.80	0.162

Table 4. Aggregated point-and-click data for Fitts' law analysis of new system

D(mm)	W(mm)	ID(bits)	MT(s)	IP=ID/MT(bit/s)
180	8.5	4.47	16.34	0.274
180	12.5	3.94	15.13	0.261
180	17	3.53	12.75	0.277
180	22	3.20	12.63	0.253

The linear regression equation derived from the results of Table 3 was MT = -0.623 + 6.148*ID, r = 0.924, F(1, 2) = 24.3, p < 0.0015. The linear regression equation derived from the results of Table 4 was MT = 2.03 + 3.22*ID, r = 0.931, F(1, 2) = 27.0, p < 0.0012. The IP value for the old system was 0.16 bit/s, while the IP value for the new system was 0.31 bit/s. Fig. 4 shows the linear regression plots for both systems.

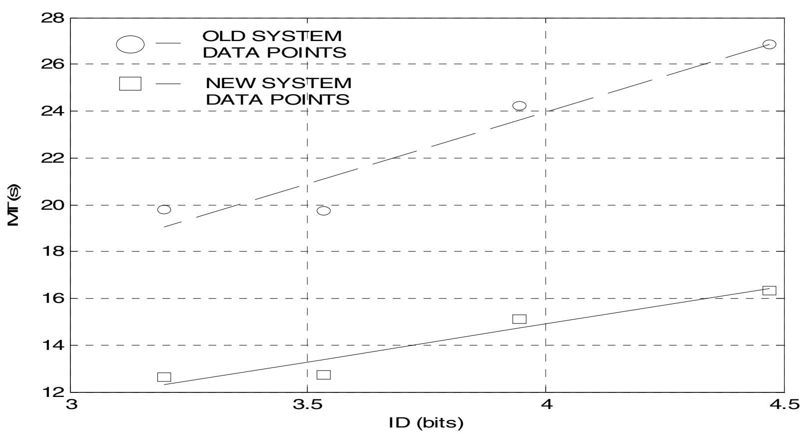


Fig. 4. Fitts' law regression lines for both cursor control systems

5 Conclusion

The results show that mean point-and-click task times are 8.45 s faster for the new system when compared to the old system. The ANOVA results prove that this difference is statistically significant. Therefore, we conclude that the new system provides faster point-and-click operations when compared to the old system.

The Fitts' law data shows that the IP value for the new algorithm is larger than that of the old one (0.31 bit/s compared to 0.16 bit/s). Also the linear regression equations and the corresponding plot indicate that the new system performs 6 - 10 s faster than the old system for a task of a given index of difficulty. One can conclude from these analyses that the new system exhibited a shorter mean object selection time, and allowed a more efficient processing of user input information (larger IP) for enhanced hands-free interaction with the GUI of a personal computer.

Acknowledgements

This work was sponsored by NSF grants IIS-0308155, CNS-0520811, HRD-0317692 and CNS-0426125.

References

1. Adjouadi, M., Sesin, A., Ayala, M., Cabrerizo, M.: Remote Eye Gaze Tracking System as a Computer Interface for Persons with Severe Motor Disability. In: Miesenberger, K. et al. (eds.): Computers Helping People with Special Needs. Lecture Notes in Computer Science, Vol. 3118. Springer-Verlag, (2004) 761-769
2. Barreto, A.B., Scargle, S.D. and Adjouadi, M.: A Practical EMG-based Human-Computer Interface for Users with Motor Disabilities. Journal Of Rehabilitation Research And Development, Vol. 37. (2000) 53-63
3. Barreto, A.B., Scargle, S.D., Adjouadi, M.: A Real-Time Assistive Computer Interface for Users with Motor Disabilities. SIGCAPH Newsletter, No. 64. (1999) 6-16
4. Basmajian, J.V. and Deluca, C.: Muscles Alive. Williams & Wilkins, Baltimore (1985)
5. Itou, T., Terao, M., Nagata, J. and Yoshida, M.: Mouse cursor control system using EMG. 2001 Conference Proceedings of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (2001) 1368-1369
6. Jacob, R.J.K.: Hot topics-eye-gaze computer interfaces: what you look at is what you get. Computer, Vol. 26. (1993) 65-66
7. Jacob, R.J.K.: The use of eye movements in human-computer interaction techniques: what you look at is what you get. ACM Transactions on Information Systems, Vol. 9 (1991) 152-169
8. N.S.C.I.A. Fact Sheets: Spinal Cord Injury Statistics. National Spinal Cord Injury Association (2005)
9. Yoshida, M., Itou, T. and Nagata, J.: Development of EMG controlled mouse cursor. Conference Proceedings. Second Joint EMBS-BMES Conference 2002. 24th Annual International Conference of the Engineering in Medicine and Biology Society. Annual Fall Meeting of the Biomedical Engineering Society, Vol. 3. (2002) 2436

A Tongue Based Control for Disabled People

Lotte N.S. Andreasen Struijk

Center for Sensory Motor Interaction, Aalborg University, DK-9220 Aalborg, Denmark
naja@miba.auc.dk

Abstract. Efficient input devices for computers and automatic equipment can improve the quality of life for severely disabled people. This work describes a new tongue- computer interface to be used by the disabled to access computers and new technologies. The new sensor system may be used during eating and speaking and may incorporate as many activation buttons as the number of characters in the alphabet.

1 Introduction

Disabled people, if unable to move their legs and arms, have a limited number of input sources available for a control system: Tongue movements, jaw movements, head movements, eye selection methods, blowing of air, brain signals and for those with intact language: speech. Therefore, current research has focused on control methods based on these inputs. Most of the resulting control systems are very visible, have little mobility and have undesirable drawbacks, such as neck pain in head control methods, headaches in eye control methods and interfering signals in speech control methods. In addition brain control methods need further research to reach an acceptable level of functionality, taking invasiveness, aesthetics and efficiency into account. A comparative study, comparing a tongue control method to a head control system and a rather simple mouth stick, resulted in all four severely disabled test persons preferring the tongue based control system, even though, it was not the fastest system [1], emphasizing the importance of aesthetics – the last thing many disabled persons want is to be even more different.

Tongue control methods are favorable since they are practically invisible and manageable for people with even severe disabilities. There have been different attempts to interface the tongue, including electrical contacts [2], hall element techniques [3] and pressure sensors. Further, a current commercially available tongue control systems is based on pressure sensitive buttons placed in the mouth cavity over the tongue [4]. The use of electrical contacts may not function during eating and talking. The technique with the Hall element has similar limitations. Further, the use of pressure sensitive sensors does not seem optimal, since normal speech and swallowing generates tongue-palatal pressures in the range of 20-60% of maximal achievable pressure [5,6], which poses demands on the detection threshold and therefore may increase the risk of fatigue. In addition, the use of pressure-based sensors may limit the maximal number of sensors that can be placed in the oral cavity, since the requirement of pressure increase the tongue-palatal contact area. Having only 9 control buttons the commercially available tongue touch keypad [4] far from utilizes the high selectivity in the

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 913 – 918, 2006.

movement of the tongue, which readily can pick out every single of our 32 teeth. This selectivity is in theory sufficient to select the alphabet directly on 26 buttons, although a different sensor type will have to be developed. Such a direct letter selection would bring the typing ability of a quadriplegic on the level of an intact person using one finger. This would make communication much more effective and attractive. Moreover, a variety of electric aids, including wheelchairs and neural prostheses, would be controllable with a wide range of commands from the same interface, making it suitable for environmental control systems.

Therefore this work describes a new inductive tongue computer interface (ITCI) to facilitate tongue-activated commands.

2 Methods

2.1 Theory

The detection method used in this work is based on Faraday´s law of induction for a coil, and uses variable inductance techniques. The idea is to change the inductance of an air-cored induction coil, by moving a ferro magnetic material, attached to the tongue, into the core of the coils (Fig 1):

From Faradays law the voltage drop across an inductance can be found as:

$$\varepsilon = -L\, di/dt = -\mu_0 \bullet \mu_r \bullet N^2 \bullet A/l \bullet di/dt \tag{1}$$

Where

$$L = \mu_0 \bullet \mu_r \bullet n^2 \bullet A/l \tag{2}$$

$$L = \text{inductance} \tag{3}$$

$$\mu_0 = \text{vacuum permeability} \tag{4}$$

$$\mu_r = \text{relative magnetic permeability of the core material} \tag{5}$$

$$N = \text{number of turns} \tag{6}$$

$$l = \text{is the average length of the magnetic flux path} \tag{7}$$

When only air is present as the core of the inductance, μ_r=1. As the ferromagnetic material is placed in the coil, the core becomes a combination of air and ferromagnetic material, and μ_r changes according to the magnetic permeability of the ferromagnetic material.

Applying a sine wave current, i, of constant peak-peak amplitude, a constant amplitude voltage drop, ε, is obtained across the coil L. Introduction of the ferromagnetic material into the air gap of the coil, results in an increase of ε, which stays increased, a until the material is removed. This will be utilized for activation of a command in the inductive tongue control system. The method resembles the known techniques used for displacement sensors [7].

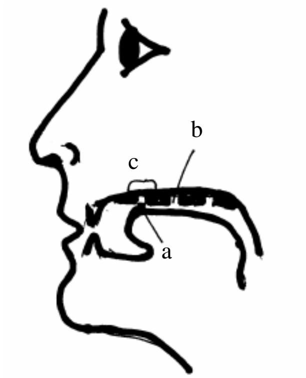


Fig. 1. The Inductive tongue control system. a: The activation unit, b: the palatal plate, c: the inductors. The tongue activates the sensors by placing the tongue-mounted activation unit at or inside a coil.

2.2 Mounting of Sensors on at the Palatal Plate

Five functional inductors were paced on a palatal plate resembling the ones used as dental retainers, see Fig.2. A silicone tube was fixed to the dental palate and carried the wires out of the mouth, see Fig. 2.

Fig. 2. Left**:** The palatal plate with inductive coils and a silicone tube leading the the wires out of the mouth. The clamps keep the plate in place at the hard palate. Right: The activation unit glued to the tongue.

2.3 Experimental Setup

The ITCS was tested in a one 37 year old, healthy female.

The palatal plate with the inductors was placed at the hard palate, and kept in place by the clamps of the plate. The activation unit was glued to the tongue using tissue glue, Fig. 2. The inductors were connected in serial, and a 30 kHz sine wave current with a 0.5 mA amplitude was applied to the inductors from a galvanically isolated current source. The subject activated the sensors by positioning the tongue in a manner that placed the activating unit in the centre of the different the coils in arbitrary order.

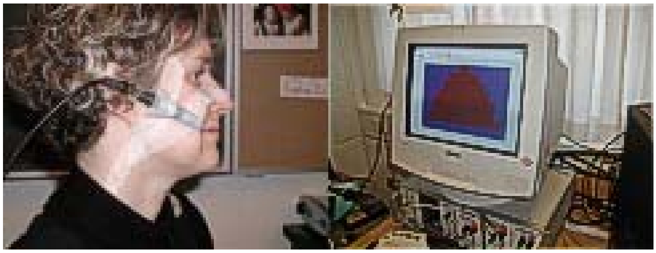


Fig. 3. Experimental set-up. The subject leads comes out from the subjects mouth. The subject is provided with a visual feedback, Fig. 4.

2.4 Signal Processing

The measured signal was amplified and rectified to obtain an envelope of the signal. Then the signal was sampled and processed using the Matlab DAQ toolbox.

Thresholding was applied to the measured signal in order to determine which sensor had been activated, and that information was fed into custom-made Matlab program providing a on-line visual display of a grid, resembling the palatal plate with the sensors, Fig. 3. From this visual display, the subject could see which sensor had been activated.

3 Results

The subject could activate desired sensors. Activation of the sensors is not conditioned the actual placement of the activation unit inside the coil, but it is sufficient to place the activation unit at the coil. As expected an activated sensor stayed activated until the activation unit was removed from that coil.

The total size of the designed inductors, given as the outer diameter was 5-6mm, indicating the possibility of having more than 25 sensors in the palatal plate of future ITCSs.

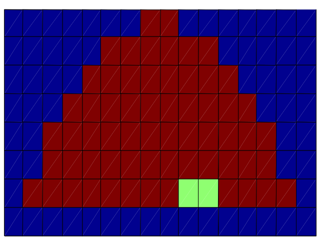


Fig. 4. Example of the visual display during activation of a sensor

4 Discussion

For a future control system to be truly successful it has, in reality to be a help for the user. This may imply that the system:

- Can be used/worn all day and night
- Is easy to use and induces a low degree of fatigue
- Is cosmetically acceptable in and outside the home of the user – preferably invisible
- Can be used to control a wide range of equipment, e.g. computers, wheelchairs, toys and prosthesis.
- Provides an efficient and quick activation of the desired function

These requirements may be met in future applications of this new Interface system, facilitating partly implantable inductive sensors.

The small size of the sensor-coils opens up for the possibility of having the whole alphabet as separate "buttons" on the palatal plate, which may lead to substantial increase in e.g. the rate of writing for quadriplegics. Future studies in tongue selectivity may reveal more information about feasibility of this possibility.

In the present work the coils were connected in serial, and the measured signal was related to all the inductances and the implicit resistance of the coil wire. A current with fixed frequency and fixed amplitude was applied, and the number of turns of each coil was the only parameter that made it possible to determinate which coil was activated. This is a very simple set-up and an advantage is that only two leads are needed to measure from and supply several sensors. Many other configurations could be considered, e.g. thresholding could be avoided using the coils as frequency detectors instead, or using several leads.

One new feature of the ITCS as compared to traditional computer control systems is that it is partly implantable, due to the activation unit. Future systems may incorporate fixation of the activation unit by piercing or by implantation right under the tongue mucosa e.g. through injection. Thus the degree of invasiveness is rather low. A questionnaire among 26 possible users of a tongue interface (laryngectomy patients in relation to a tongue controlled voice prosthesis) [8] indicated acceptance of implantation in 50% (average age: 65 years) of the cases, were indecisive in 7% of the cases and 33% (average age 70 years) would not accept implantation. The rest did not know or wanted more information. The arguments for not accepting the implant included: The asked persons felt too old, needed more information, or would prefer to have the implant during their laryngectomy surgery.

The advantages of the partly implantable sensors as compared to the traditionally pressure based tongue interfaces is, besides the possibility of having a high number of sensors/buttons, that practically no force is needed for activation and that the activation therefore is faster and less fatiguing. Further, the system is not affected by the tongue palatal-pressures related to normal eating and speaking.

Future work will focus on tongue selectivity, incorporating wireless control in the ITCS, to avoid wires coming out from the mouth and develop command strategies to control a wide range of devices.

Acknowledgements

We thank Professor Peter Svensson, for kindly being supportive with dental materials, Professor Osamu Okuno for support regarding the SUS stainless steel used for the activation unit and Helle I Mortensen for support with the experiments and the technicians at SMI for technical support. This work was made possible by grants from the Danish Technical Research Council, and Det Obelske familiefond.

References

1. Lau C. and O´Leary S., ”Comparison of Computer Interface Devices for Persons with Severe Physical Disabilities”, The American Journal of Occupational Theraphy, vol. 47, pp.1022-1030, 1993
2. Clayton C, Platts RGS, Steinberg M, and Hennequin JR, “Palatal tongue controller”, J. of Microcomputer Applications, 15, pp. 9-12, 1992
3. Buchhold N, “Apparatus for controlling peripheral devices through tongue movement, and method of processing control signals” ,United States Patent no. 5,460,186, 1995.
4. Patent: WO9307726, New Abilities Systems Inc., 1993.
5. Müller, E. et al. ”Perioral tissue mechanics during speech production”. In J. Eisenfeld & C. DeLisi (Eds.) ”Mathematics and computers in biomedical applications”. Elsevier Science publishers B.V., 1984”
6. Hayashi R, Tsuga K, Hosokawa R, et al.,”A novel handy probe for tongue pressure measurement”
7. Göpel, W, Hesse, J, Zemel, JN “Magnetic sensors”, in “ Sensors, a comprehensive Survey” Volume 5,VCH, Verlagsgesellschaft mbH, D-6940 Weinheim, 1989.
8. Nielsen Y. D. Intra-Oral Tungestyret Stemmeprotese til Laryngektomerede, Master Thesis, Department of Health Science and Technology, Aalborg University, Denmark, 2005

An Alternative Chinese Keyboard Layout Design for Single-Digit Typists

Ming-Chung Chen[1] and Ting-Fang Wu[2]

[1] National Chiayi University, Department of Special Education,
300 University Rd., Chiayi, Taiwan
mtchen@mail.ncyu.edu.tw
[2] National Taiwan Normal University, Graduate Institute of Rehabilitation Counsel
162, Sec. 1 Hoping East Rd., Taipei, Taiwan
tfwu@ntnu.edu.tw

Abstract. This study designed an alternative Chinese keyboard layout for single-digit typists and evaluated the efficacy of this innovative layout design. The new eight row by five column keyboard layout was designed based on the principles of alternative keyboard design. Eight college students with proficient keyboarding were involved in this study. The repeated measurement experimental design was used to compare the speed and accuracy of keystroke among the four different keyboard patterns: QWERTY, Alternative, Revised-QWERTY, and Random-Alternative. The experimental results indicated that the subjects' typing speed is fastest when utilizing the QWERTY layout (63.86 symbols/minute), followed by the Alternative (56.02 symbols/minute), Revised QWERTY (53.39 symbols/minute) and the Random-Alternative keyboard (49.94 symbols/minute). There is no significant difference among QWERTY, Alternative, Revised-QWERTY, and Random-Alternative layouts on the subjects' typing accuracy. The possible causes of the unpredicted results and suggestions for further studies were discussed.

1 Introduction

As the information technology advances rapidly in the everyday contexts, text entry becomes an important and prerequisite skill in the pursuit of higher educational and vocational goals, especially for people with disabilities. An appropriate text entry system can decrease the barriers of typing with computer and enhance the efficiency of the text entry [5]. The efficiency of the text entry depends on the suitability of the control interface, including proper selection methods and keyboard layouts [5,6]. The keyboard layouts have significantly impacted users' typing speed and accuracy, traveling distance of fingers, ease of learning, and efficiency of the interaction between the user and the computer. A suitable keyboard layout is essential for single-digit typists to enter text efficiently. Numerous researchers have engaged in designing English keyboard layout for the typists with upper extremity impairments [8]. LIAISON and Chubon keyboards are two of the most conventionally adopted ones [4,5,6]. LIAISON is a keyboard emulator with alternative keyboard layouts designed for individuals with high-level spinal cord injuries (SCI). The defaulted keyboard layout is designed for increasing users typing speed by minimizing the distance

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 919 – 925, 2006.

between cursor movements. The Chubon keyboard is designed to accelerate text entry for single-digit typists or typists with a typing-stick. It was designed by centralizing letters used most commonly in the English language and by placing frequently used letters together in close vicinity.

However, adaptation of the Chinese keyboard for people with disabilities has seldom been explored [2, 3, 7], with none of them focusing on the single-digit typists' needs. Proper keyboard layouts lacked, single-digit typists or typists with a typing-stick may encounter difficulties in typing Chinese efficiently. Therefore, a Chinese alternative keyboard layout must be developed for single-digit typists or typists with a typing-stick. The proposed Chinese alternative keyboard will serve as a foundation to design various keyboards, such as contracted keyboards and on-screen keyboards.

Additionally to layout design, Chinese input method is another critical issue for typists who use Chinese. Two major input systems in Chinese are the phonic-coding input method, and the other is the pattern-coding input method. Individuals who adopt the phonic-coding input method need primary spelling capability. For individuals lacking spelling ability, the pattern-coding input method may be more adaptive. Researches have demonstrated the effectiveness of Chinese pattern-coding input method [2,3]. An alternative Chinese keyboard layout was developed based on the Da-Yi input method, a pattern-coding input [2]. This alternative Chinese keyboard layout has demonstrated its effectiveness and ease of learning for individuals even with mental retardation. However, most Taiwanese are not acquainted with pattern-coding input method. Phonetic symbol systems are more popular in Taiwan since all people learned it in elementary school. Phonetic symbols are arranged in their order on the QWERTY keyboard as well as the English alphabetical keyboard. The arrangement of phonetic symbols is inefficient for single-digit typists. Therefore, this study develops an alternative Chinese keyboard layout based on the phonetic input method for single-digit typists to accelerate their typing speed and accuracy.

2 Methodology

2.1 Chinese Alternative Keyboard Layout Design

The keyboard layout was rearranged based on the principles of human computer interaction that were used for designing the alternative keyboard [4]. The major principles adapted for designing this alternative Chinese keyboard layout were: 1) *centralizing the most commonly used symbols,* 2) *adjoining the consonants and vowels that were most frequently used together in close proximity,* 3) *placing the consonants that were used as prefixes on the left side of the keyboard layout, and vowels that were used as a suffixes on the right side.*

In order to identify the most commonly used symbols in Chinese, a Chinese word database investigated by the Ministry of Education in Taiwan was used as the foundation to calculate the frequency of each Chinese character. This database investigated 1,579,771 Chinese characters that were collected from Chinese publications, and 5,063 Chinese characters were found the most commonly used.

Researchers analyzed the frequency of each phonic symbol of those 5,063 Chinese characters in advance. Based on the results of the frequency analysis of each phonic symbol and the three keyboard designing principles mentioned above, researchers redesigned a new alternative Chinese keyboard layout. This new developed keyboard is an eight column by five row layout. The placement of the keys on the alternative and the QWERTY layouts are displayed in Figure1 and Figure2. The 17 phonograms in the central area on the alternative keyboard layout contributed 58.77 % of the phonograms used for spelling the 1,579,771 Chinese characters. Conversely, the contribution of the 17 keys in the central area on the QWERTY layout was 47.05%.

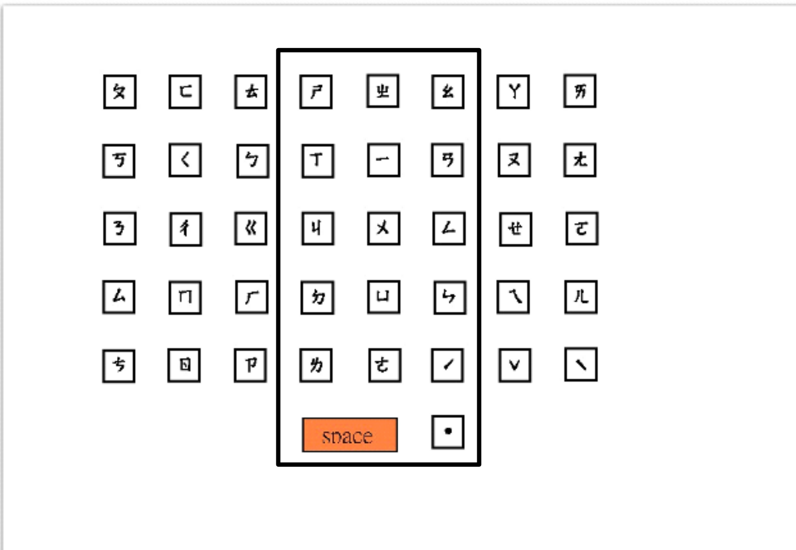


Fig. 1. Arrangements of the keys on alternative layout

2.2 Research Design

A repeated measurement experimental designed was applied to compare the typing speed and accuracy of eight subjects among the four different keyboard layouts: QWERTY, Alternative, Revised-QWERTY, and Random-Alternative.

2.3 Subjects

A convenience sample was collected by posting recruiting messages on the BBS of the college campus. The message outlined the objectives of study and stated the criteria of the participants. Only those people without perceptual or upper extremity deficiencies could volunteer and be recruited. The first eight respondents, meeting the criteria and willing to agree with the requisitions of the experiment, were selected as subjects. Eight undergraduate students, two males and six females, participated in this study. All of them were familiar with Chinese phonetic input methods and adopted it as their major input method. Participants were asked to use their index finger of their dominant hand to simulate the typing performance of single-digit typists. This study may be limited in the reliance on an entirely able-bodied sample, which may restrict generalization of the research results to the people with physical disabilities. These selected subjects could separate the effects of the disability from the effects of the keyboard layout designs [1]. A replication of this study applying subjects with disabilities may show interaction effects that may interfere the research results.

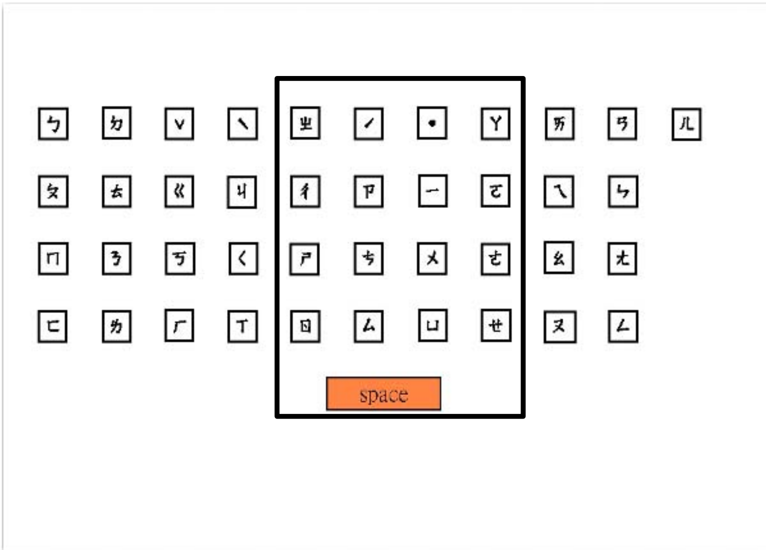


Fig. 2. Arrangements of the keys on QWERTY layout

2.4 Equipment

Researchers utilized *Editor of Unlimiter 1* to design four different keyboard layouts: QWERTY, Alternative, Revised-QWERTY, and Random-Alternative. *Unlimiter 1* is an alternative computer access device, which is similar to Intellikeys. For the purpose of controlling the participants' previous experiences of the QWERTY layout, researchers reversed the QWERTY layout vertically and horizontally to the designed a Revised-QWERTY layout. This revised layout provides a novel keyboard pattern while preserving the biomechanical relationship of the keys of QWERTY layout. The Alternative layout was developed based on the principles of the keyboard layout design. The Random-Alternative layout was designed for comparing the effect of keyboard style from the keyboard pattern. Researchers selected the phonograms in random and assigned them to the layout with eight row by five column, just like the Alternative layout.

The characteristics of typing text significantly influence typing performance [7]. Typing material which reflect real situation instead of randomly would enhance typing performance. Therefore, ten newspaper articles were selected as typing text.

The first 100 Chinese characters of each article were decoded into Phonetic symbols consecutively. A computer software program developed by Visual BASIC 6.0 was used to show the Phonetic symbols on the computer screen. The Phonetic symbol was presented in a proper order and the subjects were asked to press the corresponding key. Keys to press and time to spend were recorded on the computer software program.

An experiment was conducted in a quiet computer lab on a college campus. Participants sat in front of a personal desktop computer around the lab and could not see which layout was adopted by other participants. A Pentium-Ⅳ 1.4 G desktop with a 17in monitor was furnished with U1 and keyboard layouts during the experiment.

2.5 Procedure

Eight subjects were randomly assigned to four groups. Each group had a different sequence of four different keyboard patterns to counterbalance the effects of learning. The sequence of the first group's subjects was QWERTY, Alternative, Revised-QWERTY, and Random-Alternative. The second group's subjects started with Alternative keyboard. The subjects used each keyboard pattern for a week. This week, the subject was asked to test the same keyboard pattern three sessions a day, four days a week. Consequently, each subject generated 12 typing records for each keyboard pattern.

Test text for each session was randomly selected from the ten articles to counterbalance the interference of experimental material. The subjects were asked to type the same text for all four keyboard patterns. The typing speed and accuracy were record by the computer software program. The process was repeated for each session until the subjects finished all the tests session of the four keyboard patterns.

2.6 Data Analysis

The data was analyzed by SPSS 10.0. Repeated sample ANOVA was used to compare whether the speed and accuracy of keystrokes of four keyboard layouts were significantly different.

3 Results

3.1 Keystroke Speed

The results of ANOVA analysis indicated that the subjects speed of applying four layouts is different ($F=11.31$, $P<.01$). The effect of posteriori comparisons found that subjects typing speed is fastest when using the QWERTY layout (63.86 symbols/minute), followed by the Alternative layout (56.02 symbols/minute), the Revised-QWERTY layout (53.391 symbols/minute) and the Random-Alternative layout (49.941 symbols/minute). Although Alternative layout is better than Random-Alternative layout, Alternative and Revised-QWERTY layouts do not significantly differ.

3.2 Keystrokes Accuracy

QWERTY, Alternative, Revised-QWERTY, and Random-Alternative layouts did not significantly differ in keystroke accuracy. This finding may be due to the high accuracy rates of subjects' performance. Accuracy of the four layouts was Alternative (99.42%), QWERTY (99.22%), Revised-QWERTY (99.18%), and Random-Alternative (99.23%).

4 Discussion and Conclusion

The subjects' typing speed is fastest when applying QWERTY layout (63.86 symbols/minute), followed by the Alternative layout (56.02 symbols/minute), Revised-QWERTY keyboard layout (53.39 symbols/minute) and Random-Alternative keyboard layout (49.94 symbols/minute). While the subjects' keystroke speed of the Alternative keyboard was slower than the QWERTY keyboard. This may be because all subjects were acquainted with the pattern of QWERTY keyboard layout. The previous experience of typing skills may influence the study results.

There is no significant difference between the Alternative keyboard and the Revised-QWERTY keyboard in subjects' typing speed. The Revised-QWERTY keyboard retains the biomechanical relationships of the QWERTY keyboard while eliminating the effect of subjects' previous experience [1]. This adaptation permits fair comparison of the alternative keyboard. In current study, each participant has only 12 typing records for each keyboard patterns. The subjects' typing speed may not reach a proficient level since they are not able to familiar with the Alternative keyboard and the Revised-QWERTY keyboard layouts in such a limited time period. Further study may extend the experimental period to allow the subjects reach the maximal performance in typing speed.

The subjects' typing speed of the Alternative keyboard design was faster than the Random-Alternative keyboard. The typing performance of the random keyboard is worse. Thus, the results of this work support that the Alternative keyboard layout is more suitable for single-digit typists.

No significant differences were found among the QWERTY, Alternative, Revised-QWERTY, and Random-Alternative layouts in keystroke accuracy. The accuracy rates of all four keyboards were above 99%. No significance was found because of ceiling effects. Further studies should heighten the difficulties of the typing task to reduce the effects of the ceiling.

This study employs healthy subjects to offer a comparison of keyboard mechanics independent of the disability effects. This method lessens the possible interaction between disabilities and keyboard layout. Thus, this study should be extended to include persons who are single-digit typists secondary to disability to ascertain the benefits of the alternative keyboard are reserved when the person has a motion restriction.

This study uses a repeated measurement experimental designed to compare the typing speed and accuracy of eight subjects among the four different keyboard layouts. A further study may use a single subject research design to allow subjects to reach the maximal performance of each keyboard pattern. The single subject research design may reduce the interference of subjects' prior experiences of certain keyboards.

Finally, this study provides an effective means of comparing several keyboard layout patterns. An addition study to compare physical keyboard versus on-screen keyboard, applying the similar methodology may generate a body of knowledge that would provide useful information when providing alternative input methods for persons with disabilities.

Acknowledgment

The authors specially appreciated Dr. Tien Yu Li's kindly suggestions. Besides, the authors would like to thank the National Science Council of the Republic of China for financially supporting this research under Contract No. NSC 91-2614-S-415-001-.

References

1. Anson, D., George, S., Gallup, R., Shea, B., and Vetter, R.: Efficiency of the Chubon versus the QWERTY keyboard. Journal of Assistive Technology, (2001) 13, 40-45
2. Chen, M. C.: The study of teaching individuals with mental retardation to learn Da-Yi input method. Bulletin of Special Education, (2000) 19, 195-214
3. Chen, M. C., Wang, H. P., and Wang-Chen, L. C.: Research on teaching Da-Yi Chinese keyboarding by using adaptive input interface._Proceedings of International Conference on Computers in Education, Taipei, (2000) 8, 535-541
4. Chubon, R. A., and Hester, M. R.: An enhanced standard computer keyboard system for single-finger and typing-stick typing. Journal of Rehabilitation Research and Development, (1988)25(4), 17-24
5. Cook, A. M., and Hussey, S. M. (Eds.): Assistive technology: Principles and practice. (2nd Eds.). Baltimore: Mosby. (2002)
6. Hurlburt, M., and Ottenbacher, K. J.: An examination of direct selection typing rate and accuracy for persons with high spinal cord injury using QWERTY and default on-screen keyboards. Journal of Rehabilitation Research and Development, (1992) 29(4), 54-63
7. Hsieh, M. C., and Luo, C. H.: Morse code text typing training of a teenager with cerebral palsy using a six-switch Morse keyboard. Technology and Disability, (1999) 10(3), 169-173.
8. Kroemer, K. H. E.: Keyboards and keying: An annotated bibliography of the literature from 1878 to 1999. Universal Access in the Information Society, (2001) 1, 99-160
9. Shih, K. K., and Lin, C. C.: A quantitative model for designing keyboard layout. Perceptual and Motor Skill, (1999) 88, 113-125

An Integrated Design for a Myoelectrically-Based Writing Module for a Controlled Prosthesis

Andres Herrera, Malek Adjouadi, and Melvin Ayala

Department of Electrical & Computer Engineering
Florida International University
10555 W. Flagler Street, Miami FL 33174, U.S.A.
{Andres.Herrera, Adjouadi, Melvin.Ayala}@fiu.edu
http//www.cate.fiu.edu

Abstract. The objective of this research was the design and implementation of a writing module that is integrated with a myoelectrical-based gripper as a potential prosthetic device that could help amputees recover some of their writing abilities. The developed module would hence offer increased functionality to current prostheses. This novel device required multidisciplinary design in mechanical, electrical, and software areas. The robust integration of these key technical areas ensured a reliable module that allowed converting voluntary muscle contractions from remaining muscles into written characters. The writing module spanned over a specific writing area using a mechanical finger to yield a realistic and feasible design. The electrical implementation involved capturing and processing real time myoelectric signals (EMG). The software section utilized an assembler based algorithm to control the overall device using the processed signals. After processing a serially inputted code, the implemented writing module accurately selected and generated the requested characters with reliable and acceptable printing quality.

1 Introduction

Advancements in prostheses technology have been achieved and continue to strive towards prosthetics, which provide greater functionality, performance, durability, and cosmetic appearance [1-3]. The use of highly advanced materials and electronics implemented into the prostheses has been limited by affordability of the technologies. We must also keep in mind the difficulty of copying the functionality of perfect extremities, such as the hand [4]. It is almost an impossible task to replicate nature, but a close resemblance to it could be obtained with current technologies.

The task of writing, as fundamental as it may seem, and often taken for granted, is as a matter of fact a challenging communication endeavor that can be appreciated only if one attempts to automate it in machines. It is a fundamental skill for young children, adolescents, and adults. Therefore, lack of development in such an essential task results in an unduly constrained interaction of the amputee with society [5]. Amputees are often challenged by many jobs that rely heavily on writing as a form of communication in general. Even with current available writing aids the amputee is

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 926 – 934, 2006.

faced with limitations of the equipment that needs to be mobilized to accomplish the writing motor task and the time that it takes to set up these devices.

It is with the intent to address and redress some of these limitations, that the this research endeavor seeks to create a writing module that will allow greater dexterity and portability to the amputees when it is attached to current upper limb prosthetic devices. The focus was thus placed in the design and deployment of a basic writing module that is designed to translate muscular activity into written characters. This module can provide, by artificial means, the acceptable implementation of the writing process.

A fundamental problem with current upper limb prosthetic devices is that they only provide minimal functionality to their users. Even though, some give articulated wrist and elbows, most prosthesis capabilities are limited to grasping and gripping objects [6-9]. Unfortunately, neither major dexterity nor fine movement has been mimicked by these devices in activities such as writing. However, this capability has not been accomplished for lack of interest, but because of fiscal constraints. One must keep in mind that the limiting factors to the development of such technologies are the funding for research and affordability of the technology to its end users [1,10,11].

2 Myoelectrically-Based Writing Module Implementation

2.1 Mechanical Implementation

The mechanical implementation of the writing module constituted one of the key components on this integral design. Mechanical design sets the basis for the electrical and software implementations because the intrinsic nature of writing is to generate physical characters. The mechanical writing system must reproduce a fairly long chain of characters that are aligned and accurate. The reproduction of a short sequence or misaligned characters could create confusion and misunderstanding. After considering writing devices, such as automatic signing machines, Dot-Matrix and bubble Jet portable printers; and while encountering numerous problems with the utilizations of such devices, the need to create a more basic model was established. Integrating the module to a myoelectrically controlled device was a challenge to the mechanical design. The spatial translation requirements for the device were greatly constricted by the integration to a prosthetic device.

A rail-actuated module is proposed as shown in Figure 1 in order to resolve the writing device implementation. The actuation mechanism uses a screw device to slide the writing tip along the rail generated axis. The strength exerted by the actuator is linear and constant. This turned out to be advantageous for writing the characters on paper because the writing tip could maintain solid contact with the writing surface. The need for error software compensation in the stroke was eliminated through accurate rail alignment. Therefore, it made this approach more computationally appealing and cost-effective with the limited resources of the microcontroller and software implementation, considered here as minimal requirements.

After proving the functionality of the mechanical design working along side with its electrical and software counter parts, the need to design a writing tip became apparent. A writing appendix needed to be placed in the mechanical system to write the characters and to accommodate changes in the altitude, texture, and overall writing surface. A device was designed to accommodate where a character ends and a new one begins. This device would avoid undesired lines connecting characters. It will also guarantee close contact to the writing surface.

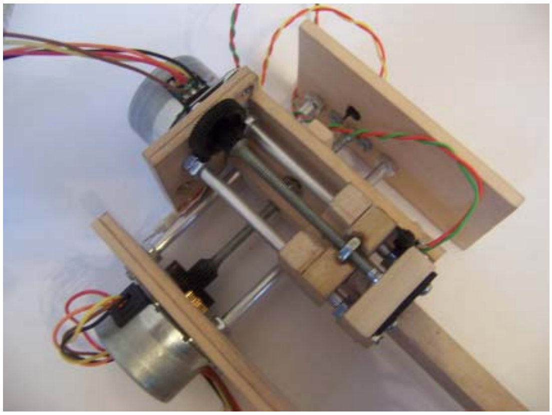


Fig. 1. Complete rail actuated system showing perpendicular rails, motors, and sensors

2.2 Electrical Implementation

After the mechanical implementation of the module was completed, an integration or adaptation to a regular electrical control system was made [6]. The electrical system captured the real time electromyogram (EMG), bio-potentials of the muscles. After the acquisition of the EMG signals, they are processed into useful data for the microcontroller. The microcontroller followed that process and analyzed the inputted data. It outputted instructions to the motors; either to write a character or to open and close the gripper for grasping task. The basic components of the electrical system are as illustrated in Figure 2.

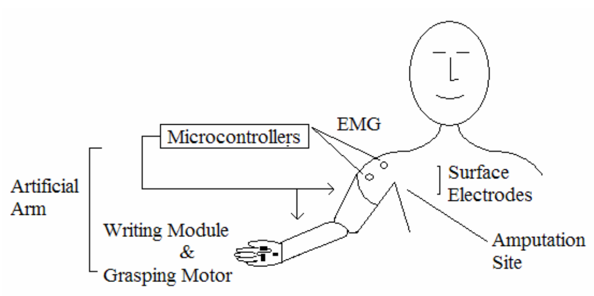


Fig. 2. Basic components of the electrical system

This module utilizes a capture system similar to those used in current prostheses as exemplified in [6-8], but it had an added electrode to distinguish between the writing and the grasping task. It has been discovered from previous experiments that adding an extra electrode connected to the ground of the main circuit would stabilize the system and the overall capturing device. This important addition prevents undesired offsets in the signals or electrical interference that includes noise in the environment, light sources, and microwave signals [6].

The electrodes are thus connected to the surface of the skin to measure the bio-potential difference among the targeted muscle groups. A set of instrumentation amplifiers were connected to the electrodes to reject common mode signals between them, and to amplify differential mode in the bio-potential measurements. After the signals have been differentiated, amplified, and rectified by the designed electrical system, they can be used as control signals. These communicate EMG activity in the muscles and differentiate which of the two EMG capture mechanism perceives more bio-potentials. These allow the microprocessor to know when EMG activity is present in the muscles and perform accordingly the bit selection, either a 0 or a 1 for the serial input.

The microcontroller board provides all the features that are needed for the implementation of the writing module. It has a serial port that can be used during the developing period, but it can also be used by the amputee to perform upgrades and updates to the firmware running in the system. The amputee would have the ability of customizing their writing module for everyday world use. The user could input detailed information, such as name and address in the system and associate them with simple codes. Therefore, the amputee would avoid the need to input a code for each character. Moreover, since the microcontroller has an 8 bit bus, up to 256 different characters can be stored and distinguished in the writing module. Higher-end microcontrollers can of course increase the number of these characters but with an additional cost. The only limitation would then be the storing capabilities of the module, which could be easily solved by adding external memory to this device.

A mosfet switching bank was used to activate the motor windings and provide isolation from the microcontroller to the stepper motors. This isolation prevented overdriving the programmable integrated circuit, thus increasing the reliability of the system. For this application two unipolar permanent magnet stepper motors were used with 7.5 degrees of step angle, which belong to a 35 mm outer diameter model series. The stepper motors included with the writing module are used in printers, scanners, and with some precision applications for their step angle. The writing module benefited considerably from the high resolution provided with the small angle step of the motors.

The system circuitry was implemented in a printed circuit board (PCB). Extra grounding was added to the circuit board to avoid interference from external noise sources. Even though, the common mode rejection ratio of the instrumentation amplifiers was very high. The circuit still received some interference in the differential inputs, mainly due to motion artifact, location and distance of the electrodes to the reference point. The PCB was fabricated in a one layer FR4 board using 5% tolerance resistors, 10% tolerance capacitor, and high precision potentiometers. The implementation of the circuit proved to be reliable by using these tolerances. Therefore, they would make the implementation of the device more affordable to the amputee during possible commercialization. Figure 3 shows the complete circuit as it was implemented.

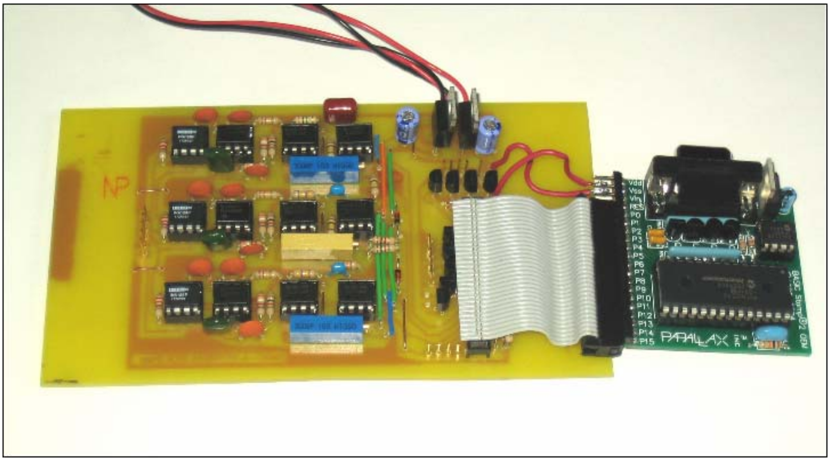


Fig. 3. Overall System Circuitry

2.3 Software Implementation

A boot routine was implemented to guarantee the correct location for the writing tip in the device. At the microprocessor start up, the boot routine would displace the mechanical system to a location, an origin, where two sensors will indicate the reached position. Once the origin was reached, the mechanical system would displace itself a specific offset to accommodate for characters. Then, the microcontroller would be ready for instructions from the myoelectric signals. Figure 4 depicts the event flow as it is being processed by the program code.

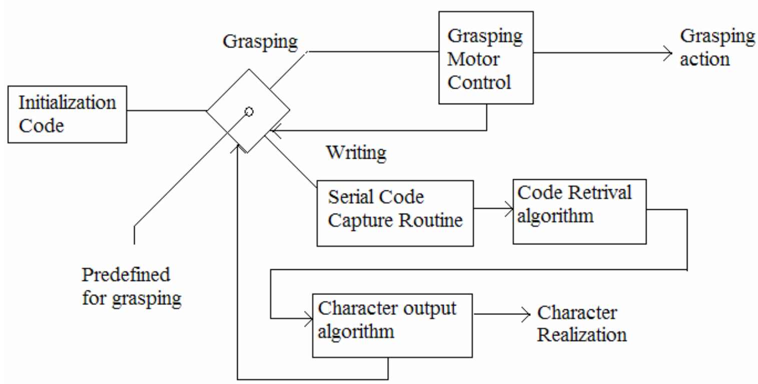


Fig. 4. Flowchart of the embedded computer program

The writing module assembler code was written in an infinite loop. The software would constantly check for bio-potential input for grasping task. The introduction of this infinite loop for grasping was a must to guarantee functionality on demand. The writing task would also be accounted for in every cycle of the loop, but only addressed when needed.

The utilization of coding schemes was viewed as the most feasible option for the user to be able to input a request for a character by using gross movements of the remaining muscle in the amputated area. The user would be able to enter a binary code that can easily be learned. The distinction between ON and OFF positions will relate directly to myoelectric bio-potentials.

3 Simulation and Testing

An incremental test (in difficulty) was assessed on the writing unit and on the modeled characters and strokes. The first logical step was to simulate a seven segment display. This pattern represented the simplest way to obtain a character. This system would use straight strokes to represent each section of the characters. However, the characterization of the letter while using this method was not as simple as activating different segments. In this case, each stroke had to be carefully planned to have the writing tip go over itself numerous time until obtaining a clear result. Figure 5 shows the first character obtained from the writing module using a seven segment display approximation.

Fig. 5. First character written by the module

The first character was obtained using 800 steps of the motor per segment. Each stroke was 3/8 of an inch long. This combination results in an approximate total resolution of 2100 steps per inch. This resolution can be enhanced to obtain a wide variety of delicate and intricate characters. As this character was generated, the need to obtain more details and number of characters grew. The next experiment consisted of generating the word CATE, the name of our Center, as shown in Figure 6.

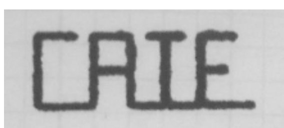


Fig. 6. Word "CATE" written not using retractable electromagnetic writing tip

The results from the imprint of the characters were not perfect, but were promising. The words were legible while using the seven segment display method. However, the issue with the undesired (connecting) lines crossing the writing plane came to the picture, literally. Therefore, the implementation of the retractable writing tip was necessary to accomplish writing characters with no running lines between them. The retractable writing tip was implemented as it was described in the mechanical design section. It was implemented using at first thick marker, but the exerted force on the electromagnetic actuator was too great to handle. Therefore, a sharpie was switched in its place to perform the writing task. The assembler code was corrected as well as the implementation of the retraction of the writing device. The code would need to be activated during the traces of a character and deactivated during the transition from the end of a character to the beginning of the next one. This ensured the disappearance of undesired lines running from one character to the other as illustrated in Fig. 7.

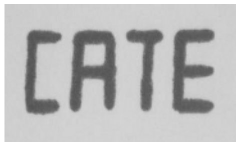


Fig. 7. Word "CATE" written using retractable electromagnetic writing tip

A further testing stage needed to be completed to determine robustness of the code, in order to obtain an accurate description as to how to trace curved and angular lines, and to study font characterization. The next goal was to trace curved lines. With this intent, drawing a letter with curves was the next landmark to attain. In this case, the seven segment display was not used for obvious reasons. The curves in the letter look promising, but it was noted that the friction on the tip of the writing device played an important role in the printing of the characters over a writing surface. Further study on the effect of the friction is needed to fully achieve the full potential of the writing device. The combination of the set of strokes in Figure 8 can generated a wide variety of font styles and characters. It will just be a matter of preparing the preferred characters for the amputee.

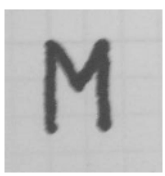


Fig. 8. Letter M resulting from combination of strokes

4 Conclusion

This research produced the groundwork for the development and future integration of a writing module for current prostheses. This device could be utilized to markedly enhance the upper limb prosthetics dexterity. Writing is one of the most delicate and dexterous tasks that human beings perform in their daily life. The designed module has provided the basis for continuous development of writing aids for amputees to be embedded in their prosthetic devices. It presented a reliable implementation on the areas of mechanical, electrical, and software engineering for embedded writing devices. However, the need for better mechanical systems, polymers, and materials played a decisive role on flaws seen in the printed characters. Especially in not being able to take full advantage of the resolution (2100 steps per inch) due to limitations in this basic mechanical design. However, an acceptable level of accuracy was accomplished, especially after introducing the electromagnetic retractable tip for the writing module.

The introduction of geared micro-stepper motors, precisely trimmed mechanics, and lightweight polymers would increase the accuracy and detail of the printed characters. Implementation of a feedback system as a positioning tool would allow the localization of the stepper motor to accurately reflect its real position since these motors can sometimes miss a step and with no proper feedback knowledge of such mishap, appropriate corrective measures can not be taken. Since limited memory size presented problems in the implementation of more characters, the introduction of a bigger size memory bank is recommended. It would allow the implementation of a large quantity of characters, words, and signatures. Commercialization of the device should proceed only through approval of the FDA, allowing as a consequence for a comprehensive study on its practical benefits and its drawbacks. The integration of the writing module to the gripper design, while seeking dependability and practicability of the device should be investigated before a prototype is constructed.

Acknowledgements

The authors gratefully appreciate the support from the National Science Foundation under grants EIA-9906600,HRD-0317692, and CNS 42615, the Office of Naval Research through grant N00014-99-1-0952, and Miami Children's Hospital.

References

1. Riddle, G. C.: Amputees & devotees. Grant C.Riddle. New York: Irvington Publishers ; Sunnyvale, CA (1989)
2. Friedmann, L.W.: The psychological rehabilitation of the amputee. Springfield, IL:Thomas (1978)
3. Cook, A. M.: Electrically-Controlled Upper Extremity Prosthetic Devices. Electronic devices for rehabilitation. John Webster, editor. New York: Wiley. Chapter 12, (1985)
4. Arem, A.: In our hands: A hand surgeon's tales of the body's most exquisite instrument. New York: Times Books, 1st Ed. (2002)
5. Vanderheim, P. Writing Aids. Electronic devices for rehabilitation. John G. Webster, editor. New York : Wiley. Chapter 10 (1985)
6. Herrera, A., Bernal, A., D. Isaza, D., and Adjouadi, M.: "Design of an Electrical Prosthetic Gripper using EMG and Linear Motion Approach". Florida Conference on the Recent Advances in Robotics (FCRAR-2004) Orlando-Florida, 2004
7. Kumar, V., Rahman, T., and Krovi V.: Assistive devices for people with motor disabilities. Wiley Encyclopedia for Electrical and Electronic Engineers. Pfeiffer (1997)
8. Muzumdar, A.: Powered upper limb prostheses: Control, Implementation and Clinical Application. Berlin; New York: Springer-Verlag (2004)

9. Scott, R.: Biomedical Engineering in Upper-Extremity prosthesis. Comprehensive management of the upper-limb amputee. D.J. Atkins, R.H. Meier III, editors. New York : Springer-Verlag. Chapter 16 (1989).
10. M. Adjouadi, A. Sesin, M. Ayala, and M. Cabrerizo, "Remote Eye Gaze Tracking System as a Computer Interface for Persons with Severe Motor Disability", pp. 761-769, Springer-Verlag Lecture Notes in Computer Science, LNCS 3118, K. Miesenberger et al., Eds., 1191 pages, July 2004; ISBN: 3540223347.
11. A. Barreto, S. Scargle and M. Adjouadi, "A Practical EMG-Based Human-Computer Interface for Users with Motor Disabilities", Journal of Rehabilitation Research and Development, Vol. 37, No. 1, pp. 53-63, Jan-Feb 2000.

Development of a Power Assisted Handrail – Handrail Trajectory and Standing Up Motion

Yoshiyuki Takahashi[1], Osamu Nitta[2], Shigeru Okikawa[3], and Takashi Komeda[3]

[1] Toyo University, Faculty of Human Life Design, Oka 48-1, Asaka-shi, 351-8510 Saitama, Japan
y-takahashi@toyonet.toyo.ac.jp
[2] Tokyo Metropolitan University, Faculty of Health Sciences, Higashi-Ogu 7-2-10, Arakawa-ku 116-8551 Tokyo, Japan
nittaosm@post.metro-hs.ac.jp
[3] Shibaura Institute of Technology, Faculty of Systems Engineering, Fukasaku 307, Minuma-ku, Saitama-shi 337-8570 Saitama, Japan
{m406022@sic., komeda@se.}shibaura-it.ac.jp

Abstract. We developed a handrail which can provide functional assistance. The result of basic experiment shows the force on the handrail could be used to expect the center of gravity of the user. And also, the handrail could assist the moving ability of a subject who cannot stand up even if using the normal handrail. In future, establish the human standing up modeling and adaptive handrail controlling for the handrail.

Keywords: Power assistance, Hand rail, Standing up.

1 Introduction

Moving is one of the important activities of daily living. For the people who have some problem on their lower limb, the handrail is useful tool for standing up, walking with support and transfer to/from the bed, wheel chair and so on. On the other hand, the handrail is difficult to adjust its location after installation. For example, the user's ability is changed due to aging or progress of a disease, the position of the handrail would be required to change to more suitable position. It will need reconstruction of the facility. And also, position of the handrail is not suitable for all people. Because, every people have different body size and level of ability. Unsuitable setting handrail could be a cause of accidents.

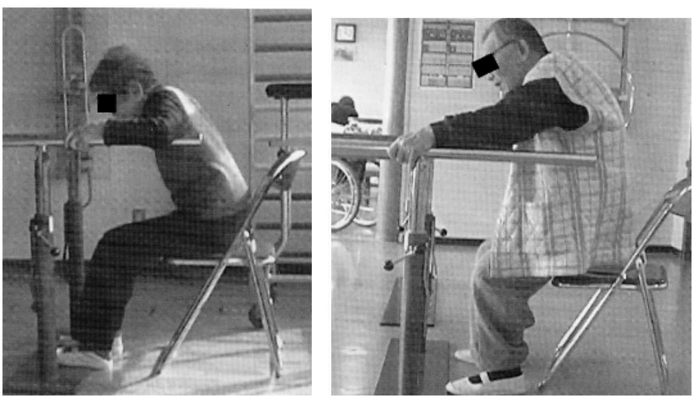
Previous researches about handrail are mainly focused on how to fix the height and shape of the handrail [1,2]. How handrail can be fully utilized to meet individuals' different physical conditions and needs has not fully been studied. In this study, we developed a handrail which can provide functional assistance. The handrail can assist the moving ability of a subject who uses the handrail to stand up.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 935 – 942, 2006.

2 Analysis of the Motor Movements of Standing Up Elderly People

We observed and analyzed the motor movements of standing up elderly people prior to the handrail development [3]. In formulating the development of the handrail, it is imperative, therefore, to consider and have a sufficient understanding of the differences among individual user's motor function. With that in mind, we intended to develop a handrail that can be induced to meet an individual user's operational patterns of standing up and is able to provide the appropriate requested support. This intention was derived from the two following factors observed in the research. First, those elderly with low motor function will exhibit little upper trunk movement (Fig. 1 (a)). Second, cases where the elderly cannot execute the center of gravity movement smoothly are very noticeable (Fig. 1 (b)). This would cause a difficulty when standing up from a chair, because an individual would normally move the center of gravity from behind the feet to the feet. To compensate for this poor motor function, it became apparent among the subjects that they would more likely to draw on the handrail towards them. To meet those needs, we thought it was necessary for the provide assistance in propelling the center of gravity forward. Based on the above-mentioned research results, we assumed that we were able to develop "Assistance type handrail" by which an individual's motor function can be fully maximized.

It was also decided that the automatic mechanism of the handrail would response accordingly to the user's operational speed and load applied to it. The handrail is motor-operated, and it is assumed that the motor is computer controlled.



(a) (b)

Fig. 1. Observed typical standing up motion of elderly subjects, (a): Leaned the handrail, pushed it down and lift her trunk, (b): Held fast the handrail, pulled it and lift his trunk.

3 Design of the Power Assisted Handrail

We designed the prototype power assisted handrail for the experiments on the assumption that the movement of the handrail could be selected freely. Fig. 2 shows

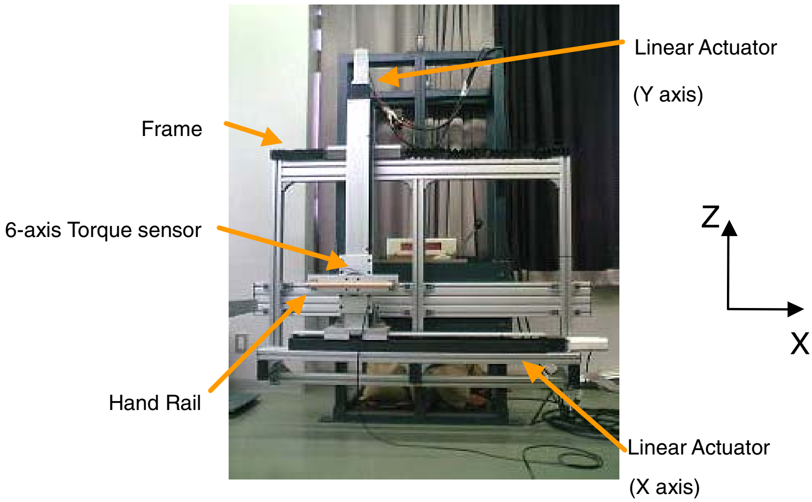


Fig. 2. External view of the power assisted handrail

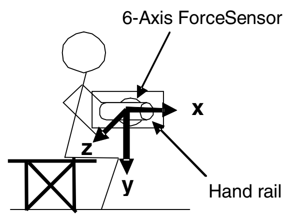


Fig. 3. Coordination system of the 6-axis force sensor

the external view of developed handrail system. Two 650 mm stroke AC servo driven liner actuators were squarely combined, and the handrail was installed at the intersection of these actuators, and the movement of the handrail was assumed to be computer controlled.

The prototype handrail developed at this time can move freely on a 650 mm by 650 mm plane. It is designed to move the 490 N load with maximum velocity 250 mm/s. Support frame of the power assisted handrail is attached to the lifter. It allows adjusting the height of this power assisted handrail system. The 6-axis force sensor (100M40A, JR3 Inc.) is attached on the base of the handrail. Fig. 3 shows the coordination system of the force sensor.

Fig. 4 shows the user interface of the handrail system control program. The handrail can be moved according to the preset trajectories and manual command. This system can store maximum 5 different trajectories. The user can direct the start position, end position and control points of the Bezier curve. Then, the control program calculates the trajectory of the handrail. Calculated trajectory is shown in upper right of the control panel and the user can understand how the handrail will move visually. The force on the handrail can be used to the trigger of the handrail movement. The direction and magnitude of trigger force is set on the control panel.

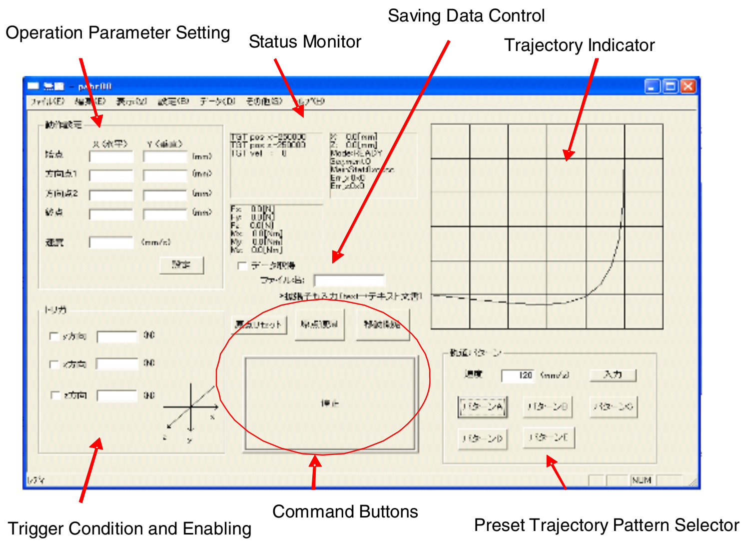


Fig. 4. Screenshot of the user interface of the power assisted handrail control program

4 Basic Experiment

4.1 Materials and Methods

To confirm the basic function and gathering standing up motion data, preliminary experiment was carried out. In this experiment, 5 young healthy subjects (4 males and one female, 19 to 20y) and two Parkinson subjects were enrolled. Young healthy subjects have no history of central nerve system disorder. Parkinson subjects could not stand up by themselves (more than Yahr III).

The handrail system was synchronized to 3D motions capture system (Vicon370, Oxford metrics Ltd.) and measured subject's joints position, handrail position and the force on the handrail simultaneously. Fig. 5 shows the structure of the experimental system. Seat level was set to the height of popliteus when the subject standing.

Fig. 6 shows the trajectory patterns in the experiment. Each grid indicates 10 mm by 10 mm. The handrail moves from lower left to upper right on the figure. Subjects sit down lower left and standing up. Left hand was used to hold the handrail. This motion is according to the previous research about changing the center of gravity when standing up [3]. The handrail started to move when the user pull the handrail more than 20 N.

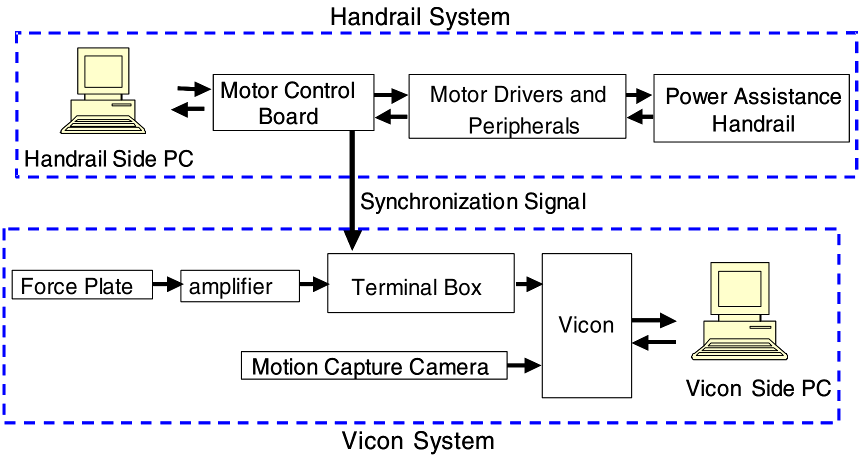


Fig. 5. Structure of the experimental system

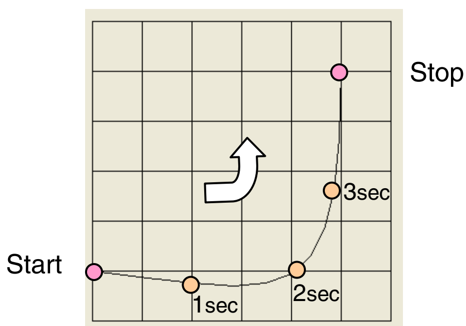


Fig. 6. Handrail trajectory pattern in this experiment

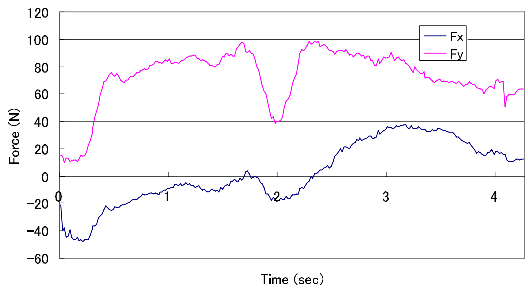
4.2 Results

Fig. 7 shows the force on the handrail of typical young healthy subject (male, 19y) and Parkinson subject (male, 73y). Time is from started to move the handrail to stop it.

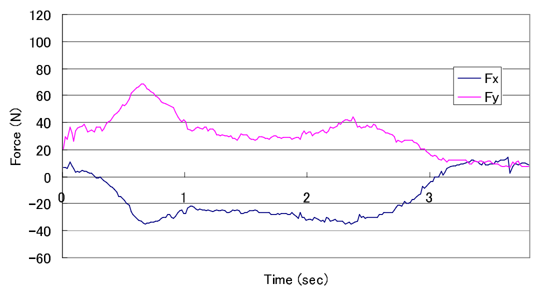
Fx is a longitudinal direction force on the handrail (Coordination system is shown in Fig. 3). Positive value means pushing the handrail to the forward and negative value means pulling the handrail to the backward. In case of young subject, the handrail was pulled initially and after 2.3 sec it was changed to pushing. At that time, the handrail moving direction was changed from horizontal to vertical. In case of Parkinson subject, pulling force was measured during the handrail moving forward.

Fy is a vertical direction force on the handrail. Positive value means pushing down the handrail. In case of young subject, the handrail was pushed down. Around two second, the force was reduced in the moment. In case of Parkinson subject, it was pushed down during the handrail moving.

When standing up using power assisted handrail, center of gravity (C.G.) of young healthy subject could be following the handrail motion. Initially, he pulled and pushed down the handrail. After that, he did not pull the handrail and pushed the handrail after 2.3 sec. This pushing force was not due to standing up motion. He was already standing up at that time. In case of Parkinson subject, he pulled and pushed down the handrail continuously. It means his C.G. could be behind the handrail during the handrail moving.



(a)

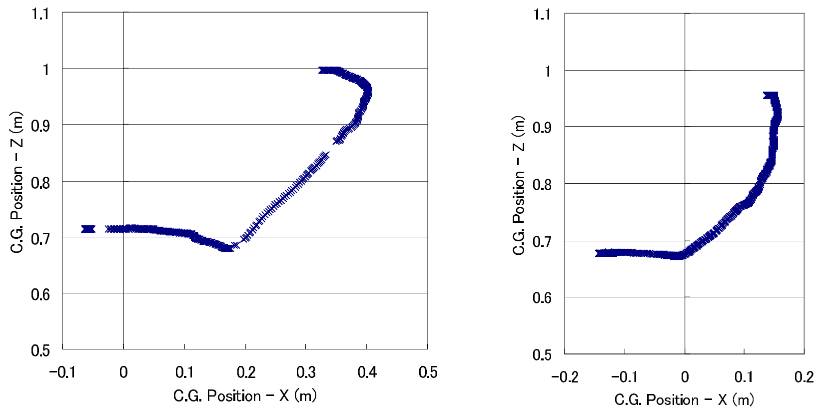


(b)

Fig. 7. Force on the handrail, (a): Young healthy subject, (b): Parkinson subject

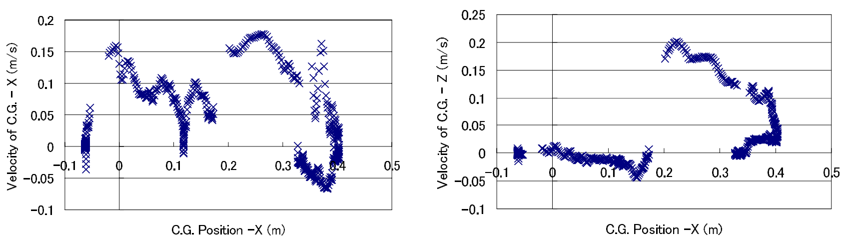
Trajectories of C.G are shown in Fig. 8. Several blank parts were unable to measuring and calculation the C.G. by 3D motion capture system. Young healthy subject's C.G. trajectory was confirmed similar movement with the handrail. C.G position was slightly going down during standing up. However, Parkinson subject's C.G. trajectory was shorter than young healthy subject in longitudinal axis and it was not down. It could be he cannot control his C.G. in longitudinal direction well.

Velocities of C.G. by young healthy subject are shown in Fig. 9 and by Parkinson subject are shown in Fig. 10. In these figures, horizontal scale is C.G. position on X axis. C.G. velocities of Parkinson subject were slower than young healthy subject's one. However, it is confirmed that the peak of velocity was found in Fig. 10 (a). In that area, he moved his C.G. quickly. It could be indicated that difficulty of keeping balance in changing posture of Parkinson subject.



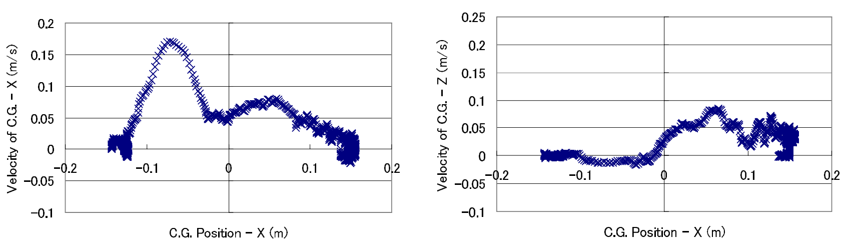
(a) (b)

Fig. 8. Trajectory of C.G., (a): Young healthy subject, (b): Parkinson subject



(a) (b)

Fig. 9. Velocity of the C.G. of young healthy subject, (a): X axis, (b): Z axis



(a) (b)

Fig. 10. Velocity of the C.G. of Parkinson subject, (a): X axis, (b): Z axis

4.3 Discussions

In this experiment, Parkinson subject was not able to stand up by himself even if using the normal static handrail. However, he could stand up by using power assisted handrail. It indicated that this handrail system could be useful tool for assisting the standing up motion of disabled people.

When standing up, we try to move C.G. forward and placed on the foot. In case of Parkinson subject, it is difficult to move C.G. forward and control well. This handrail system led his changing C.G. to forward and support his trunk. Establish the human standing up model, finding suitable trajectory, velocity control and C.G. estimation will be next step of this research.

5 Conclusion

We developed a handrail which can provide functional assistance. The result of basic experiment shows the force on the handrail could be used to expect the center of gravity of the user. And also, the handrail could assist the moving ability of a subject who cannot stand up even if using the normal handrail. In future, establish the human standing up model and develop an adaptive handrail controlling for the handrail.

References

1. Tatsuo O., Akiko I., Yoshihiro E.: Research of a taste of "diameter and height" of handrail, Proceedings of Architectural Institute of Japan (2000) 1031-1032 (Japanese)
2. Tatsuki T., Yoshitaka N., Hidehiko H.: Study of Design of Handrail, Proceedings of the Japan Society of Mechanical Engineers, Bioengineering Division (2005) 111-112 (Japanese)
3. Osamu N., Masaya A.: The Relationship between Handrail Grasping Conditions and Grip Strengths in the Aged, Rigakuryoho Kagaku 17(3). Tokyo (2002) 175-179 (Japanese)
4. Osamu N., Kazuo N., Takako T., Hiroshi T.: Power Assistance Type Handrail Which Synchronizes with User's Standing Up Operation, Proceedings of the 3rd World Congress of the International Society of Physical and Rehabilitation Medicine –ISPRM (2005) 477-2

Trial Usage of Rehabilitation System: Simple Driving Simulator for the Driving Skill Evaluation of People with Cerebrovascular Disease: CVD

Kaoru Inoue[1], Yuko Ito[1], Yumi Ikeda[1], Atsuko Tanimura[1], Keisuke Suzuki[2], Yoshiyuki Takahashi[3], Tadashi Suzuki[4], Madoka Fujimoto[5], Masahiro Tsuchijima[6], and Takashi Komeda[7]

[1] Tokyo Metropolitan University, 7-2-10 Higashiogu, Arakawa-ku,Tokyo, 116-8551, Japan
inoue@post.metro-hs.ac.jp
http://www.tmu.ac.jp
[2] Hibari-en, Health Care Facility for the elderly requiring long-term care, 1277 Ikkui Showa-cho Nakakoma-gun Yamanashi 409-3852, Japan
[3] The University of Tokyo Center for Collaborative Research, 4-6-1 Komaba, Meguro-ku, Tokyo, 153-8904, Japan
[4] Tokyo Teshin Hospital, 2-14-23 Fujimi, Chiyoda-ku, Tokyo, 102-8798, Japan
[5] Tamba City Aogaki Clinic, 114-1 Sawano, Aogaki-cyo, Tanba, Japan
[6] Nagoya City Rehabilitation and Sports Center, 1-2 Mikanyama, Yatomi-cho, Mizuho-ku, Nagoya, Japan
[7] Shibaura Institute of Technology, 307 Fukasaku Minuma-ku, Saitama, 330-8570, Japan

Abstract. We have developed a simple tentative version of the driving simulator for rehabilitation using a Personal Computer (SDS). We described the over-view of the system and the trial usage of this system for people with Cerebrovascular disease. The test results of the subjects whom an occupational therapist in attendance assessed to be able to drive a car showed values almost the same as the data of the healthy subjects. The effectiveness of the system was suggested in the preliminary experiments and a case study.

Keywords: Cerebrobascular disease, Rehabilitation, Driving simulator.

1 Introduction

We developed several rehabilitation systems to provide quantitative data; commencing with upper-limb rehabilitation systems using virtual-reality technology; haptic devices [1,2,3]. We think that the symptoms of patients that we can show by quantitative data has to show them with a numerical index, we therefore accomplished some projects.

There are many reports on driving simulators [4,5,6]. It seems that subjects of these studies are patients with obstructive sleep apnea syndrome (OSAS), neurologi-cally impaired persons, and elderly people. There are a lot of study reports to prove the therapeutic effects of using a driving simulator. The investigation on effects of simulator-based driver training showed the possibility of neurologic recovery [5]. We also showed that driving skills of elderly persons were restored by training using a driving simulator [7].

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 943 – 949, 2006.

The authors made a simple driving simulator, and investigated the trial experiments for people with cerebrovascular disease.

An Automobile is an important transportation device for elderly people and handicapped persons in Japan. Therapists are often asked to give opinions concerning the driving ability of elderly people and handicapped persons. However there are few devices available with which to determine their ability. In Japan many precision driving simulators have been developed, but those are very expensive, and large. Those require specialized knowledge for the operator. So except for large rehabilitation hospitals and centers, many small hospitals, and welfare institutions for elderly people can't buy them. We thought we need simple driving simulators for small medical or welfare institutions, in addition to precision driving simulators. Especially, the driving skills of people with Cerebrovascular disease (CVD) often came into question among medical staff. Recently many hospitals and welfare institutions have met the needs of the times and adopted electronic medical chart systems. So transition from an analog to digital data was expected.

We developed a simple tentative version of the driving simulator for rehabilitation (SDS) using a Personal Computer. Using this system, therapists would be able to offer good rehabilitation services of increased quality. It is hoped that this system can be applied to develop effective rehabilitation techniques, to motivate clients by arousing their interest, and to quantitatively verify therapeutic results.

2 Materials and Methods

2.1 The Driving Simulator: SDS

This system consists of PC, liquid crystal monitor, and control device, Microsoft Sidewinder Force Feedback. We used software of Visual C++, and made an application system for the skill evaluation, and training of driving. Test items are classified into reaction time of the upper limb, steering operation, and accelerator/brakes pedal. This report introduces reaction time measurement, and basic steering operation test. The test of reaction time consists of two types and the basic steering test, four types.

Fig. 1. The SDS System

2.2 Reaction Time Tests

In this Task the subject is expected to push a button immediately after he discovers either a static or dynamic target (12colors, 4 patterns) unless an "X" mark is presented on a monitor. The system records the reaction time of the subject between the button push and the display of a target on a screen as well as the button he pushes. As for the interval of target appearance, an arbitrary setting is indicated. The system is so set that the circles appear at random intervals, from one second to three seconds.

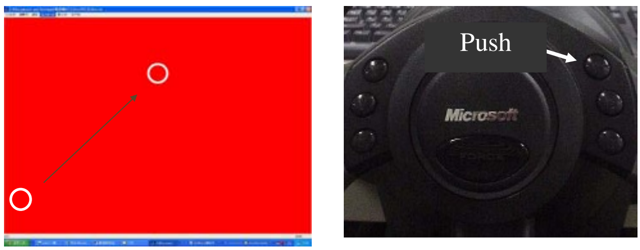


Fig. 2. The RT-Tests

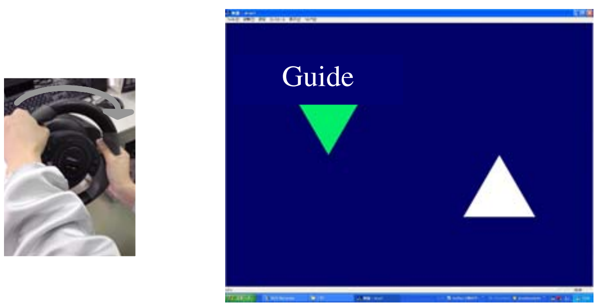


Fig. 3. The STR-Tests

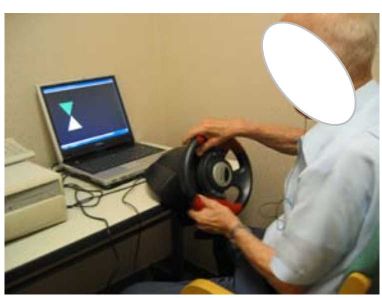


Fig. 4. A subject operating the SDS

2.3 Basic Steering Wheel Operation Tests: STR-Tests

The upper green triangle guide moves from side to side on a screen. In this task, the subject moves the white triangle precisely towards the triangle guide by operating a steering wheel. There are four movement guide patterns; STR-L (low speed), STR-M (medium speed), STR-H (High speed) and STR-R (random speed; the guide shows random movement).

3 The Experiment

3.1 Purpose

We carried out preliminary experiments of the SDS. The data of healthy subjects will then be used as reference data to evaluate the degree of disability of clients.

The aims of the experiments are 2 points. One is to get data on subjects with CVD, whose driving ability were to be kept. The second is to confirm the effectiveness of the SDS regarding rehabilitation service for a client.

3.2 Method

Tests carried out were the RT Test, and STR-Test of the SDS. The same tests were given to both the healthy subjects and to people with CVD. The subjects with CVD operated the device of the SDS with the non-plegia side hand. About steering operation test (STR-Test), the people who able to use both hands operated a steering by two-handed. The static analyses were done to study the correlations among all test results. We used "SPSS" Ver.13.0 for statistics analysis.

3.3 Results and Discussion

Tests were carried out from Dec.12th 2003 Aug.31st 2005.

The healthy subjects were 51 people aged 20 through 73. The results indicate the average of the measured values. The over sixty subjects showed remarkable individual differences. Among elderly persons, some showed results similar to the young group, on the other hand, there were those who needed about twice the time of the young (Figure4,5,6). The subjects with brain injury were 6 people aged 62 through 85.

6 people were with Cerebrovascular disease; CVD, 3 people were right hemiplegia, The other were left hemiplegia. They could walk with canes, do self-care, and had no or slight higher brain dysfunction. They were not patients at a hospital, lived at home in their community, and needed to drive their car in their daily life. An occupational therapist in attendance supposed that they might drive a car. Providing against an emergency, the experiment for subjects with CVD was carried out in a clinic in their community. All subjects agreed to this study.

On our recent findings, it was suggested that the RT Test had relationships with mental functions of subjects [7,8]. On healthy subjects, RT-Tests and STR-tests showed the tendency, the young group received a better grade than elderly people.

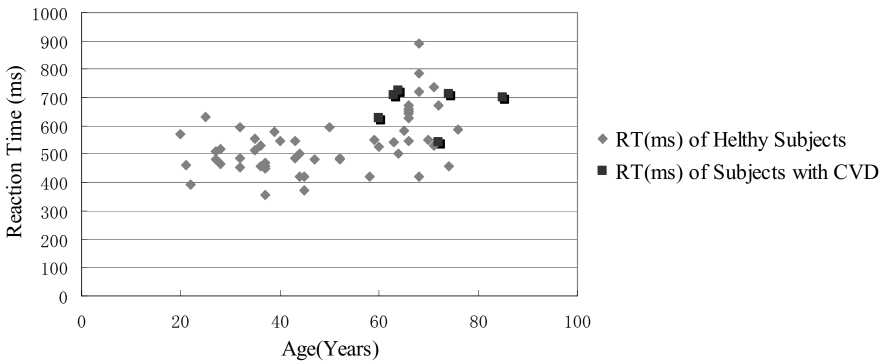


Fig. 5. Age and the Results of Reaction Time Tests

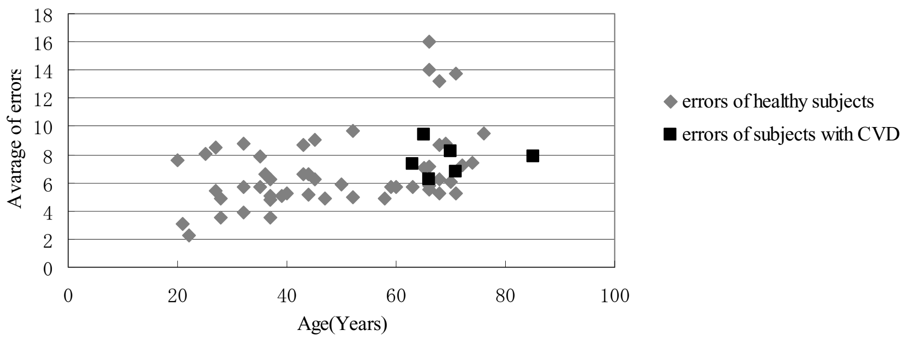


Fig. 6. Age and the errors of STR-Tests (STR-H)

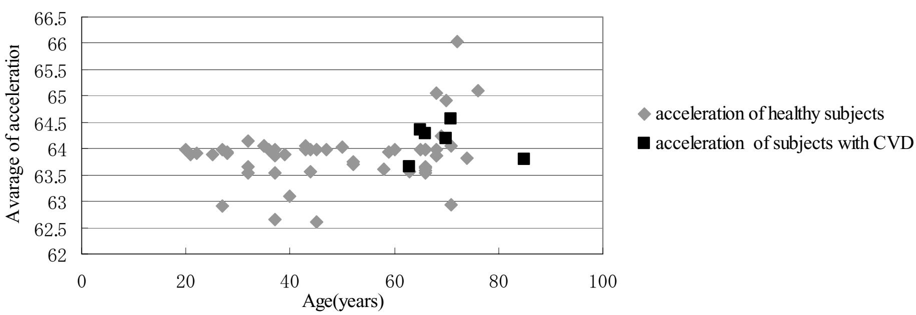


Fig. 7. Age and the acceleration of STR-Tests (STR-H)

On the RT-Test, the average of healthy subjects of 60's was 626.4ms, the standard deviation was 121.4ms. The results of subjects with CVD were 629.4ms, 707.9ms, 724.7ms with ages. The average of healthy subjects of 70's was 589.0ms, the standard deviation was 99.9ms. The results of subjects with CVD were 542.5ms, 711.5ms with ages. We did not get healthy data of 80's. But the 85-year-old man with right-hemipregia showed good result as the results of 60-70's healthy persons. On the errors and acceleration of the STR-Test, as well as the RT-Test, the results of some subjects with CVD showed at the same level as the results of the healthy subjects.

We thought these results suggested the possibility that they could drive a car if they satisfied other conditions.

About the data of people with higher brain dysfunction, or dementia, we thought there was still room for further study on these points. We will examine this problem at the next stage.

3.4 Case Presentation

A man with right-hemiplegia, who was one of the subjects with CVD, was able to drive a car by rehabilitation. Now he is enjoying driving a car everyday.

He is a 64-year-old right-handed Japanese male with a three-year history of CVD, he came to the clinic because of the decline of his strength, and hoping to get physical exercise. He could walk using a cane, do self-care. He was living with his mother and his wife. His occupational history was a bus driver, and the manager of an automobile repair shop. In his living area, there are few supermarkets, restaurants, clinics, and public institutions. He had to drive a car to live. While he could not drive a car, he and his family had to use a taxi, but taxi fare was expensive for him. His mother and his wife did not have driving licenses. Japanese elderly women do not have driving licenses by Japanese cultural influence. So the Japanese elderly man has a cultural role to take his family to and from by car. Many Japanese men are proud of this role.

But he could not have confidence to drive a car with his handicap. An occupational therapist in attendance fed back the experiment results to him among progress of rehabilitation. An occupational therapist judged that he had the ability to drive a car, and suggested to him to take tests of the SDS.

He told that "the results of the experiences and experiment of daily training gave me confidence ".

He proceeded with his hobby, and could fulfill his role as driver for his family. We realized that quantitative data developed motivation for the client. We thought the independence of car driving contributed to improvement of his and his family's Quality of life; QOL.

4 Conclusion

We developed a simple tentative version of the driving simulator for rehabilitation using a Personal Computer (SDS). The characteristics of the SDS were shown, and the effectiveness of the system was suggested in these preliminary experiments.

The results of some subjects with CVD showed skills at the same level as with the results of the healthy subjects. At that stage, we collected data of people with CVD, to analyze. We thought it was the effective way to show the digital data to a client in order to carry out rehabilitation smoothly.

These results would serve as basic data, when therapists try to evaluate and treat clients who need rehabilitation in driving skills. We would like to continue developing of useful devices for rehabilitation in the future.

Acknowledgements

We were able to carry out this study funded by the Ministry of Education, Culture, Sports, Science and Technology, and by the Mitsui Sumitomo welfare foundation. We would like to show our thanks for the support. Last, but not least, we would like to thank volunteers.

References

1. Y. Takahashi, T. Terada, et. Al.: 'Upper-Limb Rehabilitation System Using Haptic Device with Basic Motion Training Program', Proceedings of Annual International Conference of the IEEE Engineering in Medicine and Biology Society, (2003):141-146
2. Yoshiyuki Takahashi, Takafumi Terada, Kaoru Inoue, Yuko Ito, Hokyoo Lee and Takashi Komeda: 'Upper-Limb Rehabilitation Exercises Using Haptic Device System', International Journal of Human-friendly Welfare Robotic Systems, (2003),14(2): 18-22.
3. K. Inoue, Y. Takahashi, et.al.: 'Developing a Haptic Device System for Upper-Limb and Cognitive Rehabilitation': An Electromyographic Anlysis, 2nd World Congress of the International Society and Rehabilitation Medicine-ISPRM, (2003): 695-698.
4. Orth M, Duchna HW, Leidag M, et.al.: Driving simulator neuropsychological [corrected] testing in OSAS before and under CPAP therapy, Eur Respir J, (2005);27(1): 242
5. Akinwuntan AE, De Weerdt W, Feys H, et.al.: Effect of simulator training on driving after stroke: a randomized controlled trial, Neurology, (2005);65(6): 843-850
6. Kotterba S, Widdig W, Brylak S.et.al.: Driving after cerebral ischemia—a driving simulator investigation, Wien Med Wochenschr. (2005): Aug; 155(15-16): 348-353
7. K. Inoue, Y. Ikeda, Y. Ito, et.al.: Development of Simple Driving Simulator for the Driving Skill Evaluation of the Elderly and the Handicapped, AAATE2005 proceedings, (2005): 219-223
8. K. Inoue, Y. Ito, Y. Ikeda, et.al.: Trial Usage of Rehabilitation System; Simple Driving Simulator for the Driving Skill Evaluation of Elderly Persons and Handicapped People, EMBEC2005 proceedings, (2005): 1950

Eye Movement Communication Control System Based on EOG and Voluntary Eye Blink

Junichi Hori[1,2], Koji Sakano[3],
Michio Miyakawa[1,2], and Yoshiaki Saitoh[1]

[1] Department of Biocybernetics
[2] Center for Transdisciplinary Research
[3] Graduate School of Science and Technology,
Niigata University, 8050 Ikarashi-2, Niigata 950-2181, Japan
hori@bc.niigata-u.ac.jp
http://www.mbe.bc.niigata-u.ac.jp/~hori

Abstract. A communication support interface controlled by eye movements and voluntary eye blink has been developed for disabled individuals with motor paralysis who cannot speak. Horizontal and vertical electro-oculograms were measured using two surface electrodes attached above and beside the dominant eye and referring to an earlobe electrode and amplified with AC-coupling in order to reduce the unnecessary drift. Four directional cursor movements —up, down, right, and left— and one selected operation were realized by logically combining the two detected channel signals based on threshold settings specific to the individual. Letter input experiments were conducted on a virtual screen keyboard. As a result, operatablility, accuracy, and processing speed were improved using our method.

1 Introduction

Developmentally disabled individuals with motor paralysis have difficulty in conveying their intentions, since the motor neurons influencing voluntary muscles are affected. In terminal amyotrophic lateral sclerosis (ALS) patients, the eye movement muscles are usually not affected and thus, using them has the potential to improve the performance of input operations in communication support systems. Several methods have been proposed that use electro-oculograms (EOGs) occurring as a result of eye movements [1,2,3,4,5]. In a simplified switching system, an automatic scanning selection must be used for information input [1]. Devices with five or more electrodes have been proposed for detecting eye movements in four directions [2,3]. However, the processing speed of these systems was slow because the selected operation mode was the eye resting for a few seconds. An eye gazing method that detects any point where the eye gazes on the screen required additional processing to reduce the drift [5].

The purpose of the present study is to develop a real-time EOG communication control system that offers improved operation and simple setup combined with high performance. We propose a device that outputs five kinds of

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 950–953, 2006.

intentions: movement in four directions (up, down, right, and left) and one selection, by transforming two channel AC-coupled EOG signals detected by three electrodes. Eye movements operate cursor movements in four directions and a selection is indicated by a voluntary eye blink.

2 Methods

2.1 Composition of EOG Communication System

In order to construct a truly useful communication support system, we made three improvements on the traditional EOG system.

First, a smaller number of electrodes should be better for practical use in daily life. By considering the usability, operatability and reproducibility, preliminary experiments determined optimum electrode positions for two channels derived by monopole electrodes of the upper part of the eye on the forehead and at the side of the eye on the temple with reference to an earlobe electrode on the dominant eye side.

Second, the detected EOG signals are amplified by AC instead of DC coupling in order to reject the drift which occurs in long-term measurements. The signals are bad-pass filtered and analog-to-digital converted. These signals are combined logically according to the threshold level and the setup time.

Third, potential changes resulting from eye movements that move a screen keyboard cursor are detected and generate four kinds of intended information by logically combining these potential changes. Moreover, the potential change due to voluntary eye blink is interpreted as selecting a certain position, similar to clicking a mouse button for achieving accuracy and input speed.

2.2 Parameter Setting

The thresholds and the setup time of EOG signals should be used for distinguishing the intentional signals from artifacts such as eye blink and ordinary saccade eye movements. The upper and lower threshold bounds in two EOG signals should be determined beforehand to identify the eye condition. The setup time that is a time interval between the first and second peaks of EOG is also an important parameter for improving the performance of the communication system. The eye movement of looking away from the certain viewpoint tends to be very slow in ordinary cases. In addition, the shape of EOG measured during involuntary eye blink looks sharper than voluntary eye movements. Considering these characteristics, the subject was instructed to move his/her eyes within 1s. After a series of trials, the optimum values of the five parameters that are maximizing the accuracy and minimizing the error were searched to fit specific individuals.

3 Experimental Results

Experiments were performed using a virtual screen keyboard to examine the usability of the proposed EOG system. The subject sat in front of a projector screen that displayed a virtual screen keyboard. The sampling frequency was set at 200 Hz. The low-cut and high-cut frequencies of the band-pass filter were 0.53 Hz (time constant of 0.5 s) and 5 Hz, respectively. First, automatic parameter setting using computer instructions was carried out to adjust the communications system. The protocol for computer instruction was a four-fold iteration, that is, the sequence of Up, Down, Right, and Left movements followed by Select. After adjusting the system, the subject was instructed to attempt to generate a sentence comprising twelve characters "GOOD MORNING" (12 letters) and 26 alphabetical letters × 3 (78 letters). Fatigue of each subject was also evaluated by a questionnaire subjectively and flicker test objectively.

Table 1 shows the experimental results of virtual letter input. The results for accuracy and the processing speed over 5 subjects were 94.1 % and 8.3 letters/min for 12 letters experiments and 87.2 % and 8.7 letters/min for 78 letters experiments. Subjects D and E were worse than the others since they wore contact lenses and grasses, respectively. Especially, the results of subject D might be influenced by his fatigue. Since input operations were done at their own pace, the processing speed was different for each subject. A maximum letter input speed of 11.0 letters/min. was obtained for well-trained subject in this case.

Table 1. Results of EOG communication experiments

12 letters experiments

Subject	Accuracy [%]	Error [%]	Speed [letters/min]	Fatigue [points]	Flicker [points]
A(unaided)	96.7	0.0	8.9	1	0
B(unaided)	98.3	0.0	8.9	1	1
C(unaided)	95.2	0.0	6.9	1	0
D(contact lenses)	89.4	0.0	7.5	3	3
E(grasses)	90.8	0.0	9.0	0	0
Average	94.1	0.0	8.3	1.2	0.8

78 letters experiments

Subject	Accuracy [%]	Error [%]	Speed [letters/min]	Fatigue [points]	Flicker [points]
A(unaided)	92.3	0.0	11.0	2	1
B(unaided)	93.9	0.0	10.1	1	2
C(unaided)	86.3	1.1	6.4	3	2
D(contact lenses)	83.2	0.5	7.9	5	8
E(grasses)	80.4	0.5	8.0	2	3
Average	87.2	0.4	8.7	2.6	3.2

4 Discussion

In this system, thresholds play an important role in improving system performance. When the threshold was set high, few of the detected EOG signals exceeded the threshold, resulting in poor accuracy. Moreover, increasing the EOG amplitude to exceed the setting threshold would make the burden onerous for the user. In contrast, when the threshold was set at a low level, the error rate increased due to unintended output. In particular, the potential change observed during involuntary eye blink tended to be misconstrued as Up output. For this reason, setting the threshold carefully makes it possible to increase the accuracy of the proposed system. Trial tests performed beforehand should determine thresholds before the system is applied to individuals.

If the system is reduced to simple functions such as switching systems, the accuracy will be improved. However, the processing speed will be slower than in the proposed method. According to the opinions from ALS patients and their supporting stuff, an accuracy of 80% is enough for communication using switching systems in the condition of locked-on. The processing speed of our system became much faster than that of the traditional EOG switching system of about 4.0 letters/min. More sophisticated signal processing and parameter setting will improve the accuracy, speed, and usability of the communication system.

5 Conclusion

We have developed an EOG communication support system that outputs five kinds of intention information by adopting band-pass filtering and a logical combination to interpret the action potentials of eye movements and voluntary winks. Consequently, we improved the accuracy and the processing speed as a result of adequate training of both the subject and the communication device. We believe that the proposed method has potential use in practical situations.

The authors thank Mr. Yamagishi (Niigata University) for his experimental help. This work was partly supported by Grant-in-Aid for Scientific Research, the Japanese Ministry of Education, Culture, Sports, Science and Technology, Japan and Grant for Promotion of Niigata University Research Projects.

References

1. Kate J.H. ten, Meer, P.M. van der: An Electro-ocular Switch for Communication of the Speechless. Med. Progress Through Technology, **10** (1983/1984) 135–141
2. LaCourse, J.R., Hludik, F.C. Jr.: An Eye Movement Communication-Control System for the Disabled. IEEE Trans. Biomed. Eng., **37** (1990) 1215–1220
3. Tecce, J.J., Gips, J., Olivieri, C.P., Pok, L.J., Consigio, M.R.: Eye Movement Control of Computer Functions. Int. J. Psychophysiology, **29** (1998) 319–325
4. Barea, R., Boquete, L., Mazo, M.: System for Assisted Mobility using Eye Movements Based on Electro-oculography. IEEE Trans. Rehab. Eng., **10** (2002) 209–217
5. Gips, J. and Olivieri, P.: EagleEyes: An Eye Control System for Persons with Disabilities. 11th Int. Conf. Tech. Persons Disabilities, Mar. (1996)

Eye as an Actuator

Marcela Fejtová, Jan Fejt, and Olga Štěpánková

The Gerstner Laboratory,
Czech Technical University in Prague, Faculty of Electrical Engineering
Technická 2, 166 27 Prague 6, Czech Republic
fejtovam@k333.felk.cvut.cz,
step@labe.felk.cvut.cz
http://cyber.felk.cvut.cz

Abstract. The I4Control® device is a new type of computer peripheral enabling non-contact control of a personal computer through eye (or head) movement. The solution emulates the computer mouse and consequently it gives its user direct access to any mouse controlled SW system. One of those is a SW keyboard, which is nowadays a common part of computer operating system. In the same way the system can be combined with other text entering solutions. Section 4 provides a brief report on corresponding preliminary experiments as well as description of some edutainment applications, for which I4Control® device seems to be very well suited. The very same solution can be applied for design of a control element in a computer-driven working environment where the hands of the PC user have to be engaged in another (primary) activity (i.e., during medical surgery or when handling greasy components of a complex machine).

1 New Approaches to Man-Computer Interaction

Computer mouse represents the most common computer interface. But the market offers even other solutions, e.g. voice input with limited vocabulary. Very innovative idea has been presented in 2002 by a team at Georgia Tech led by Dr. Kennedy and Dr. Bakay. They proved that it is possible to achieve even direct brain-computer interaction / their sophisticated system allowed to interpret the brain neural signals for affecting position of computer mouse. A computer system controlled by the power of human thought left science fiction and it emerged in the real world.

Simultaneously, our laboratory started its project of a computer controlled by eye movements [2]. Both teams were driven by a similar motivation: to provide new means for communication for paralysed people based on the limited abilities they have retained. Since even severely handicapped patients retain ability to move their eyes most often, we have decided to utilize this information about eye movements as an input for changing position of a computer mouse.

The idea is not novel - control of a computer by eye movements is being studied for several years already and they produce very good results. Let us review briefly some of the existing solutions. The LC Technologies Eyegaze System [1] provides an eye-controlled human-computer interface (HCI),combining a video camera mounted below the computer monitor that observes the user's eye with a specialized image

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 954 – 961, 2006.

processing software, which analyses the video images of the eye and determines its gaze-point on the monitor screen in real time. QualiEye [2], part of QualiWORLD, gives simple, efficient, hands-free control of the PC using a standard USB web cam for tracking eye movements. Electrooculography is used at Boston College in the project called EagleEyes [4] for moving cursor at the computer monitor. Great disadvantage of this method is the need to stick the surface electrodes on the user's face and to remove all artifacts originating from eye blinking or movements of face muscles. Further on, its application is limited to devoted software products that has to be used to control the PC. At the Cambridge University [5, 6] the project Dasher uses a IR camera that detects the eye position based on the contrast between the light absorbed by the pupil and the reflected light. Each of the presented systems has its advantages and draw-backs. Some of them are even commercially available. Unfortunately, most of them are very expensive and ready to provide a single function, namely control of a PC mouse. Frequently, they allow no head motion (position of head is fixed) because the camera is mounted on the PC monitor. Consequently, they set unnatural constrains on its user.

Our intention was to build a system, which is simple to operate and which can be used at home conditions. To achieve this goal, we have decided to combine experience gained by the existing systems with the advantages of current technology. We aim for design and development of a simple and inexpensive solution, which allows head movement and that ensures multiple functions depending on the software it is complemented with.

The resulting novel computer interface called I4Control® device emulates the computer mouse through eye movements. Its user gets direct access to everything, what can be controlled by the PC mouse. This is the case of a software keyboard, which is nowadays a common part of an operating system. Consequently the I4Control® user is ready to type using the software keyboard and he/she can write the text documents, send the e-mails, brows the web pages, etc. The resulting periphery enables physically handicapped users to communicate with their environment or even work or study on their own. The device can be used even in medical domain e.g. by patients with decreased ability to express themselves or by patients with cerebral palsy.

2 Design Principles for a New Solution

The I4Control® device is *the first generation of a product representing a novel application of several established technologies*, namely:

- *Videooculography* (VOG), based on recording eye movements by camera, has been in use for scanning eye movements for 30 years, approximately. The main advantage of VOG is that it is an absolutely non-invasive method, which is able to produce results of precision requested by intended applications. The quality of its results is determined by the following factors: quality of the grabbed image with observed scene representing the user and his/her eye, powers to compensate for movement of user´s head, to recognise the eye irrespective of different positions of head and to detect eye movements. A traditional VOG solution relies on an intricate system consisting of several cameras and often it requires strictly fixed position of the head of its user. This setting proved useful in laboratory environment but it is too restrictive to be used in real life setting.

- *Recent miniaturization of technologies*. The I4Control® system applies a novel original design: a single miniature camera is bound to the head of the user, e.g. it is attached on his/her spectacle frame (see Fig. 1). Now, the user´s eye position is recorded with respect to the relative system of coordinates the center of which coincides with the head of the user (and consequently, it copies all his/her movements). Usage of this "free or flying" placement of the camera eliminates most of the inherent problems of traditional approaches (there is no more need for multiple well located cameras, for head fixation and for demanding algorithms for recognition of an eye in the scene). Data registered by the camera are processed in a small device, which is independent of the controlled computer. This solution brings a significant additional value to the user who is free to move at his/her will without any limitations (stable position with fixed head is eliminated).

I4Control® interpretation of the obtained data thus becomes much simpler. Change of the eye position is evaluated as a result of finding the deviation from the idle balanced position. In fact the eye allocation has no immediate effect on the current cursor position on the monitor but it determines the direction of its movement (in the same was as the joystick). To enable usage of the system by persons with a disorder of equilibrium organ manifested as fast involuntary eye movements (nystagmus) it is necessary to introduce an insensitivity zone for detection of deviation. This zone must be adjustable individually according to the patient's handicap.

The miniature camera is connected to the PC using standard communication interface (USB) equipped with corresponding software which provides the handicapped person with the means to control the PC simply by moving his/her eye from the idle position (when the user looks directly ahead).

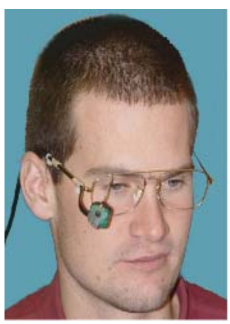


Fig. 1. Location of the camera in the system I4Control®

The system I4Control® is designed in such a way that the camera becomes another input periphery and it can be linked to any PC (see Fig. 2 for its building blocks). The system gains its full advantage as soon as it is combined with the software keyboard (see Section 4). There is no more need for specialized SW for handicapped users. On the contrary, the user gets access to any common installed SW, which he/she can control. The resulting communication module overcomes the problems of compatibility of drivers between individual versions of systems or hardware platforms, which represents serious draw back of the other gaze-based solutions.

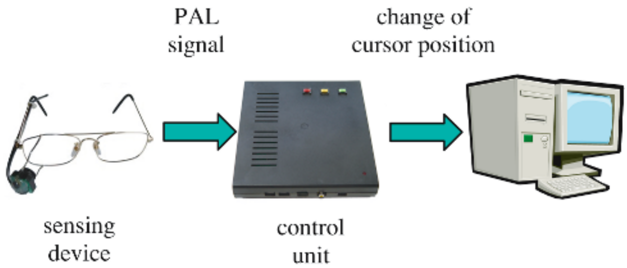


Fig. 2. Basic block structure

The Figure 2 describes the rough structure of our solution. Let us provide some more details concerning principal functions ensured by its individual building blocks:

Sensing device. The actual location of the user´s eye has no direct influence on the cursor position on the monitor but it determines the direction of its movement (as much as a joystick). The core of the system is black&white camera with a CCD sensor and discrimination 208 x 156 pixels for recording the eye movements. The output of the camera is an analog PAL signal that is digitized in the control unit.

Control unit. Principal function of the control unit is to:

- digitize the PAL signal from the black&white camera,
- detect pupil position from the digitized PAL signal and
- transform the actual pupil position into the coordinate system used for the computer mouse.

The digitizing block consists of an A/D converter with brightness feedback. Data specifying position of the pupil is evaluated by the CPU (processor) from the digitized PAL signal. The resulting information is stored in the memory for later use.

In the system I4Control® version 1.0 the SRAM 512 kB memory is used. The processor uses for the detection of pupil position the method of histogram thresholding and other types of filtration (e.g. cubic filter). Another task of the processor is to control USB bus ensuring communication with a PC. The outputs of the system are the change x- and y-coordinates of the mouse cursor.

Personal computer interprets the data obtained from the control unit in a way, which is analogous to interpretation of the input obtained from the standard computer mouse. Moreover, the computer plays an important role during the calibration process described in the next section.

3 Pilot Implementation

Current technology offers means, which can be used to implement the suggested design. The main task is to ensure correct interpretation of the image crapped by the tiny camera. For any user one has to identify his/her balanced position of the eye, first. This position is used as a reference point to which all the new data are compared – see Fig. 3, which explains how this comparison is used to control cursor movements: Direction of cursor movement is computed as the appropriate deviation from the balanced position. In the first option (a), the cursor stands still (since the eye is in the balanced position) while in the second option (b) the cursor is moving to the right.

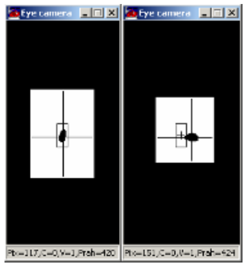


Fig. 3. PC cursor control

A necessary condition for good functioning of the suggested system is a clear video image. For that purpose it is important to ensure reasonable light conditions in the considered scene, namely good visibility of user´s eye. The simplest way how to achieve this goal is to take care of correct and high quality light conditions in the whole working environment. Since they are not difficult to ensure, the system produced according to the upper mentioned design principles has been constructed. It can work as a new computer periphery provided we offer the user possibility to ensure some additional control signals corresponding to mouse click or double-click. Eye can easily signal that by blinking: the system interprets a situation when his/her eye remains closed for a certain period (which is longer time than spontaneous eye blink) as a click. At the moment, the pilot version of the resulting device is tested in our laboratory and by several volunteers.

The main advantage of our solution is that it represents a technology, which is safe and available to everyone – it meets the demands of democratic AI. It can be produced in mass and nearly everybody can learn to use it. Several students of Jedlicka´s Institute for handicapped young people volunteered to participate in pilot testing of the I4Control® device. The device proved to be very well accepted by handicapped persons (e.g. a boy who lost both hands as a result of an accident or a boy with myopathy) who have learned to use it quickly. For example one of the users succeeded to create a simple picture in a drawing software while another one used it to navigate through an education software with relatively complex menu.

An important feature is that the used I4Control® hardware is based on easily available and financially not demanding off-the-shelf components which are assembled into a simple, mobile, efficient and fault tolerant product with good resolution which can be easily installed at home of its user. The resulting computer periphery is not the first device ensuring gaze-based man-computer interaction. Certainly, it is the first one offering following significant advantages to its users:

- it does not pose on them any unnatural constraints: the user does not have to remain fixed in a single position (so that it does not interfere with his/her usual habits),
- the device is as easy to operate as glasses (it is absolutely non-invasive – there is no need to glue special marks etc.),
- it has affordable price and it is simple to install,
- moreover, the user needs no demanding instruction.

When working with the system I4Control®, the user acquires certainty and precision after some time. The length of this period is user specific and it depends on his/her abilities to control own eye movements and patience. There are people who master the system within few minutes, other need 1-2 hours. The training period can be cut even shorter when the first steps with the system are presented in the form edutainment. The first corresponding didactic toy is briefly described in the Section 4.

4 Experiments

We have performed a number of experiments with the I4Control® system. Let us mention briefly text entering and our edutainment applications.

Text entry. The most natural way how to write a text with the I4Control® system is to utilize *software keyboard* which is currently included in common operating systems. This solution makes it possible to write the text documents, send e-mails, brows the web pages, etc. Of course, this type of text entering needs certain level of skills from the user, who has to be able to position the cursor rather precisely onto the chosen key displayed on the computer monitor – let us remember that there are at least 40 keys on the SW keyboard! But I4Control® can be easily combined with more sophisticated SW solutions, e.g. with *Dasher SW* system. We have invited several volunteers who were using I4Control® for the first time to test this combination. They were given a simple task, namely to write a sentence “I am hungry.” Without any preliminary training they were able to do so in about 12 minutes on average.

It is clear that it is not possible to gain full advantage of both mentioned approaches without certain training. To support users who want to utilize the I4Control® system for text entering immediately and without any need for preliminary training as well as the users who might face problems with precise positioning of the cursor on the computer screen we have designed a simple SW communicator. When writing a text with this communicator the user has to choose several times between 2 large panels on the screen: at the beginning the first panel displays letters **a** to **m** while the second covers the rest of the alphabet. If the user chooses the first panel in the next step the content of both panels changes and he/she is supposed to choose between one containing the letters **a** to **g** and the other with letter **h** to **m**, etc. In this way he/she reaches the intended letter upon 5 eye movements and 5 clicks. A novice user is able to write the sentence “I am hungry.” in less than 5 minutes.

The reported experiments point to some strong points and to weaknesses of the considered solutions, only. Now, it is time to ensure reliable usability study the results of which can become a basis for recommendation which solution is best for specific purpose and for each individual. We plan to apply for that purpose evaluation methodology designed within the COGAIN project [3].

Training through edutainment. To master the control of cursor position through eye movements requires certain experience and skill. The user has to learn how the system interprets his/her efforts. In this training phase, one does not need the computer itself. The I4Control® sensor and the control unit can be connected directly to the remote control of any appropriate toy and the training can start. We have used for that purpose a remotely operated crawler (caterpillar tractor) built from LEGO®. The task of the person undergoing the training is to transport the load (a ball) of the crawler into the wicket. Instead of a joystick, the user has to use for that purpose his/her eyes only. This solution proved very helpful for persons who have no prior experience with computers, even for pre-school children. Later the faculty gained when playing with toy crawler is reused for moving the picture of crawler on the screen. Now, the user understands the relation between the crawler (cursor) position and his/her eye movements, he/she learns to estimate the speed of cursor movement and the world of computer games is opened for him/her.

5 Conclusion

I4Control® device offers simple means for control of the PC to people with serious physical handicap. Using it they can better integrate into the knowledge society despite their handicap and take advantage of ICT as well as access to the latest information on the Internet. The system is small, light and sets no restrictions on its user who can move freely - it disturbs the user at the minimum level. The hardware is based on easily available and financially not demanding off-the-shelf components. It is simple, efficient, fault tolerant. In combination with the software it is flexible and has good resolution. Advantage of this system is its simple installation (it is absolutely noninvasive) and the fact that a user of a new computer interface needs no special instruction. When working with the system I4Control® for control of the cursor, the user acquires certainty and precision after some time. The length of this period differs from user to user and depends on his/her abilities to control own eye movements and patience. At present it is possible to control the PC using the system I4Control® similarly to standard PC mouse or keyboard. This type of periphery enables to physically handicapped users to work or study on their own. The device can be used even in medical domain e.g. by patients with decreased ability to express themselves or by patients with cerebral palsy.

Purpose, use and users. The I4Control® device is designed and constructed first of all for handicapped users. That is why the stress is on its simple installation, usage, robustness and low price. Since the system emulates a classical computer mouse it serves as an intermediary between its user and all mouse-controlled SW applications installed on the PC. Consequently the I4Control® user can share the wealth of all available SW products (e.g. e-learning applications, e-mail, web access) and he/she is no more forced to master and buy expensive SW complementing some existing assistive technologies. The I4Control® device is an efficient tool to ensure e-inclusion of challenged persons. Several Czech institutions for handicapped persons are planning to provide their clients with the I4Control® device.

The technical innovative *IT content* of the I4Control® device as a novel computer periphery is clear. It represents a significant contribution towards extending the acceptance and understanding of IT through societal benefits it provides to challenged persons. It mediates to its handicapped users access to information, to social/healthcare services and to education. From the long term perspective it supports their employment opportunities and social integration of its users.

Additionally, the design applied in the system can be reused in much wider context, e.g. in industry to drive a robot (or another computer controlled machinery/technology). The I4Control® device can serve well in many medical applications as a part of a diagnostic and/or a therapeutic system – the corresponding development and experiments are currently under way.

Acknowledgments

The presented research and development has been partially supported by the grants NR8261 Therapy of paresis n. abducens by means of biological feedback and IST Network of Excellence IST-2003-511598 COGAIN (Communication by Gaze Interaction).

References

1. http://www.eyegaze.com/doc/cathuniv.htm [cit. 2004-04-13]
2. http://www.qualilife.com [cit. 2004-04-13]
3. Andersen H.H.K. et al.: State of the art report of evaluation methodology, D6.1 available at http://www.cogain.org/results/reports/COGAIN/D6.1.pdf
4. Gips, J., Olivieri, P. EagleEyes: An Eye Control System for Person with Disabilities. In Proceedings 11th ICTPD, March 1996 Los Angeles, California. Published in Human-Computer Interaction: Applications and Case Studies, M.J. Smith and G. Salvendy (eds.), Elsevier, 1996, pp. 630-635
5. Cavendish Laboratory, Cambridge: The Dasher Project [online]. Last review 2004-01-07 [cit. 2004-01-20]. http://www.inference.phy.cam.ac.uk/dasher
6. Ward, D.J. Adaptive Computer Interfaces. Ph.D. thesis. Cambridge: University of Cambridge, Cavendish Laboratory, Inference Group, 2001. 144 p
7. Kšára, M.: Special text editor. Bachelor project. Prague: Czech Technical University in Prague, Faculty of Electrical. Eng., Gerstner laboratory 2006. 78 p

Nonlinear Mapping of Pupil Centre Coordinates from Image Sensor to Screen for Gaze Control Systems

Gintautas Daunys and Nerijus Ramanauskas

Siauliai University, 76353, 141 Vilniaus Str., Siauliai, Lithuania
{g.daunys, n.ramanauskas}@tf.su.lt

Abstract. Exact calibration in real time is critical for gaze control systems. Usually measurements are mapped to points on screen using coefficients obtained from calibration data. The mathematical model of pupil centre/ eye corner gaze tracking system was proposed. 6 parameters were used to describe both eyes movement on image sensor. Experimental results show good correspondence with model over all screen area. As some parameters are user specific and other can be measured independently, the number of calibration points could be reduced drastically, keeping nonlinear mapping.

1 Introduction

Disabled people may have difficulties when controlling a PC by classical peripheries. Interaction with computer using only eye movement is alternative computer control method. The precise mapping from image co-ordinates to computer monitor screen co-ordinates in real time is necessary for efficient computer control. The objective of this study is to find best mapping for pupil/eye corner gaze tracking system.

2 Related Works

A calibration procedure is required to compute the mapping between the measurements and the eye orientation. A typical calibration procedure presents the user a set of visual targets that the user has to look at while the corresponding measurement is taken. From these correspondences, a mapping or calibration function can be computed.

Morimoto et al. [1] use 9 points for calibration and a second order polynomial calibration function. The polynomial is defined as:

$$s_x = a_0 + a_1 x + a_2 y + a_3 xy + a_4 x^2 + a_5 y^2, \tag{1}$$

$$s_y = b_0 + b_1 x + b_2 y + b_3 xy + b_4 x^2 + b_5 y^2, \tag{2}$$

where (s_x, s_y) are screen coordinates and (x, y) is the pupil–corneal reflection vector. The coefficients $a_0 - a_5$ and $b_0 - b_5$ are the unknowns and can be founded using least squares.

Fitting higher than second order polynomials has increased the accuracy of the system [2], but the second order requires less calibration points and provides a good

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 962 – 965, 2006.

approximation. Simpler linear models have also been successfully used. White et al. [3] use a simple linear model with independent components, and noticed that in practice, higher order polynomial functions do not provide better calibration. Using a more complex calibration model the calibration is better near the monitor edges.

Fitting a polynomial function over the whole monitor screen is also not a requirement. For example, Zhu and Yang [4] construct a 2D linear mapping from the vector between the eye corner and the iris centre to the gaze angle. After calibration, the gaze direction is computed by interpolation. They report an average error of about 1.1° using subpixel accuracy for tracking the eye corner and iris-centre, and about 3.3° using pixel level accuracy.

Advanced remote eye gaze trackers that are being researched today basically try to eliminate two problems, the need of calibration per user session, and the large restriction on head motion. A system suggested by Morimoto et al. [5] estimates eye gaze without calibration and allowing free head motion. But it requires the calibration of the camera with respect to the monitor and light positions, and a model of the user's eye. Experimental results show an accuracy of about 3° using synthetic images. Another system described by Yoo et al. [6] uses four LEDs around the monitor screen to project these corners on the corneal surface and the fifth LED is placed near the CCD camera lens to create a bright pupil image. Using the invariance property of cross ratios under perspective, they compute the point of regard with an accuracy of about 2°. The advantage of this method is that it does not require camera calibration. Newman et al. [7] gives example of system that first computes the face pose in 3D, and then compute the eye gaze. The system runs in real time, but the accuracy is very low, about 5°.

Ohno et al. [8] have proposed a gaze tracking system, which requires only two points of calibration and the calibration is done only once per user. They introduce an eye ball model width two personal parameters: the radius of cornea curvature and the distance between the pupil and the centre of corneal curvature. These values were regarded as constant for different users, so this may result additional errors. The tracking in this system is done only for one eye.

3 Model

We analysed pupil/eye corner gaze tracking system. Its experimental setup is shown in Figure 1. The model was developed to describe eye image formation process.

X and Y axis of laboratory co-ordinate system are oriented as the same axis of computer monitor screen. Axis Z is perpendicular to monitor screen. Other coordinates system X'Y'Z' is related to image sensor, which is used in video camera. The eye rotations are described in laboratory system. We could obtain the eye image, by linear projection of eye structure to X'Y' plane of camera system.

We aligned a unit vector **e** along eye optical axis. The projections of the **e** vector to plane X'Y' is proportional to measured pupil centre coordinates. We assume that orientation system X'Y'Z' is defined by two angles. One parameter defines proportionality between end of vector **e** and real pupil movements on image sensor.

The gaze point coordinates on screen depend on distance from eye to screen. Also we must take into account that eye sight axis doesn't correspond with eye optical axis,

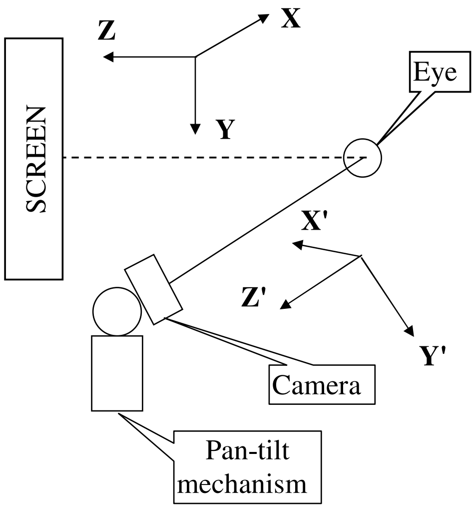


Fig. 1. Experimental setup of pupil/eye corner gaze tracking system and co-ordinates systems, for image formation

because the last doesn't go through the fovea. For every eye we have one parameter – the angle between optical and sight axis. So minimum number of needed parameters is 6.

4 Results and Discussion

The experimental results of fixation of 25 calibration points square grid is shown in Figure 2. It is evident from plot that there exist nonlinear distortions. Especially the distortions are big in upper corners of the grid.

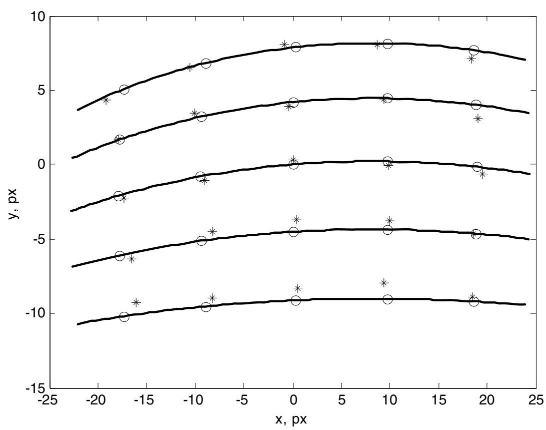


Fig. 2. Experimental results of raw coordinates of pupil centre during fixation 25 calibration points (o - simulated, * - measured)

The developed model predicts many distortions, shown in Figure 2. Remaining errors could be caused by lighting changes in tertiary eye positions. Eye cornea optical properties is other source of errors.

Developed mathematical model of system suggest nonlinear mapping using less calibration points, comparing with polynomial mapping. Because sight axis shift from optical axis is constant for user, we can save it from first carefully carried out calibration session for future sessions.

Acknowledgement. This research is supported by the Lithuanian Science and Studies Foundation (Project title: Environment Control by Gaze System for the Disabled; Registration Number: V-06003) and FP6/IST Network of Excellence project COGAIN.

References

1. Morimoto, C., Koons, D., Amir, A., Flickner, M.: Pupil detection and tracking using multiple light sources, Image Vis. Comput. 18 (4) (2000) 331–336
2. Cherif, Z., Nait-Ali, A., Motsch, J., Krebs, M.: An adaptive calibration of an infrared light device used for gaze tracking. In: Proc. of the IEEE Instrumentation and Measurement Technology Conference, Anchorage, AK, (2002) 1029–1033
3. White, K.Jr., Hutchinson, T., Carley, J.M.: Spatially dynamic calibration of an eye-tracking system. IEEE Trans. Syst. Man Cybernetics 23 (4) (1993) 1162–1168
4. Zhu, J., Yang, J.: Subpixel eye gaze tracking. In: Proc. of the 5th IEEE Int. Conf. on Automatic Face and Gesture Recognition, Washington, DC (2002) 131–136
5. Morimoto, C., Amir, A., Flickner, M.: Detecting eye position and gaze from a single camera and 2 light sources. In: Proc. of the Int. Conf. on Pattern Recognition, Quebec, Canada, (2002)
6. Yoo, D., Kim, J., Lee, B., Chung, M.: Non contact eye gaze tracking system by mapping of corneal reflections. In: Proc. of the Int. Conf. on Automatic Face and Gesture Recognition, Washington, DC, (2002) 94–99
7. Newman, R., Matsumoto, Y., Rougeaux, S., Zelinsky, A.: Real time stereo tracking for head pose and gaze estimation. In: Proc. of the Int. Conf. on Automatic Face and Gesture Recognition, Grenoble, France, (2000)
8. Ohno, T., Mukawa, N., Yoshikawa, A.: FreeGaze: A Gaze Tracking System for Everyday Gaze Interaction. In: Proc. of the symposium on ETRA (2002) 125-132

Keyboard Adaptations for Children with Cerebral Palsy

Ting-Fang Wu[1] and Ming-Chung Chen[2]

[1] Graduate Institute of Rehabilitation Counseling
National Taiwan Normal University
162, Ho-Ping E. Rd., Sec.1 Taipei, Taiwan, R.O.C.
tfwu@ntnu.edu.tw
[2] Department of Special Education, National Chiayi University
300 University Rd., Chiayi, Taiwan, R.O.C.
mtchen@mail.ncyu.edu.tw

Abstract. The purpose of this study is to systematically investigate the effects of keyboard adaptations for children with cerebral palsy. Twelve children aged from 7 to 15 years old participated in this study. Keyboard adaptation strategies were developed based on the individualized assessments. A group comparison experimental design was selected to examine the effectiveness of keyboard adaptations. Speed and accuracy of typing Chinese were compared before and after keyboard adaptations. The results indicated that children with cerebral palsy did increase their typing performance after implementing keyboard adaptation strategies. The results of this study can provide health and educational profession a reference when serving children with physical disabilities.

1 Introduction

The keyboard is the most common method of controlling computer for most of people and almost all computers equipped with a keyboard. In addition, it is generally most convenient and cheapest access method. However, children with cerebral palsy usually experience obstacles in accessing a standard keyboard because of their abnormal postures and movements. Rehabilitation professions have been involved in the application of assistive technology devices that enable individuals with disabilities access technology. How to adapting standard keyboard or providing an alternative keyboard to fit special needs for individuals with cerebral palsy is a challenging task for rehabilitation professions.

A review of the literature revealed only few studies on investigating the effects of adapting keyboard on persons with physical disabilities. McCormack reported the effects of the use of a keyguard used on typing speed and accuracy in an 8-year-old boy with cerebral palsy [1]. The results showed that the use of a keyguard increased the subject's typing accuracy, but decreased his speed. Battenberg and Merbler compared of the effect of computer input device on task performance of 40 developmentally delayed and 40 control kindergarten children [2]. Children in both groups performed better in an alphabet matching task and a spelling task when using a touch-sensitive screen compared to a traditional keyboard. Touch-sensitive screens seem to be more beneficent for children with young ages. Lau and O'Leary employed a

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 966 – 972, 2006.

descriptive case study design to compare subjects' performance using three computer input devices: the Tongue Touch Keypad, the HeadMaster, and the mouth sticks with standard keyboards [3]. Four adolescents with physical disabilities participated in this study. Two of them were cereal palsy, and the other two were clients with spinal cord injuries. Results showed input speed to be the fastest with the mouthstick, followed by the HeadMaster, and then the Tongue Touch Keypad. Accuracy of inputs did not vary significantly among three devices. Most of research employed case studies to report the effects of computer input devices. The lack of systematic measurements of subjects' improvements restricts the generalization of those studies. Therefore, the purpose of this study is to use a systematic method to investigate the effect of keyboard adaptation on children with cerebral palsy. The typing speeds and accuracy of twelve children with cerebral palsy before and after keyboard adaptation will be compared. The results of this study will provide rehabilitation and educational professions a reference when serving children with cerebral palsy.

2 Methods

2.1 Participants

Twelve children with cerebral palsy participated in this experiment with their parents' consent. Their personal characteristics are described in Table 1.

Table 1. Characteristics of the participants

Partici-pant	Age	Sex	Diagnosis	Severity	Computer use (hours/day)	Internet use (hours/day)
P1	8Y3M	Male	Quadriplegia	Severe	0.25	0.25
P2	14Y4M	Female	Diplegia	Severe	N/A	N/A
P3	7Y	Female	Athetoid	Severe	2	0.5
P4	15Y4M	Female	Diplegia	Severe	N/A	N/A
P5	8Y4M	Male	Diplegia	Severe	1	1
P6	8Y	Female	Mix type	Moderate	1	N/A
P7	12Y2M	Male	Diplegia	Moderate	0.5	0.5
P8	8Y8M	Male	Diplegia	Moderate	0.7	0.5
P9	9Y8M	Male	Hemiplegia	Moderate	0.5	0.2
P10	10Y9M	Female	Hemiplegia	Moderate	0.5	0.5
P11	8Y3M	Female	Diplegia	Moderate	1	1.5
P12	12Y2M	Female	Diplegia	Moderate	0.5	0.5

Note: N/A means the participant does not use computer regularly in their daily life.

Their ages ranged from 7 years to 15 years 4 months with an average of 10 years 3 months. All participants were diagnosed with cerebral palsy, seven of them were diplegia, two of them were hemiplegia, one is spastic quadriplegia, one is athetosis and the other one is mix typed cerebral palsy. Hours for daily computer and internet use were indicated in Table 1.

2.2 Instruments

The instruments used in this study include two self designed computer software programs "Keyboard Adaptation Assessment" (KAA) and "Chinese Keyboarding Performance Evaluation" (CKPE). KAA was used to determine appropriate keyboard adaptation strategies. The CKPE was used to record speed and accuracy of Chinese keyboarding inputs in this study.

Keyboard Adaptation Assessment (KAA). KAA was developed to assess clients' keyboard performance and provide proper recommendations of keyboard adaptations based on the results of assessments. Clients were asked to perform several keyboarding tasks, such as press a specific key or press two keys simultaneously, which were generated automatically by KAA. The results of KAA will indicated the client's keyboard access range, difficulties in keyboard operations and suggestions for individualized keyboard adaptations based on the assessment data.

Chinese Keyboarding Performance Evaluation (CKPE). A computer software program was designed to record the speed and accuracy of Chinese Keyboarding Performance. Figure 1 illustrates the design of software screen.

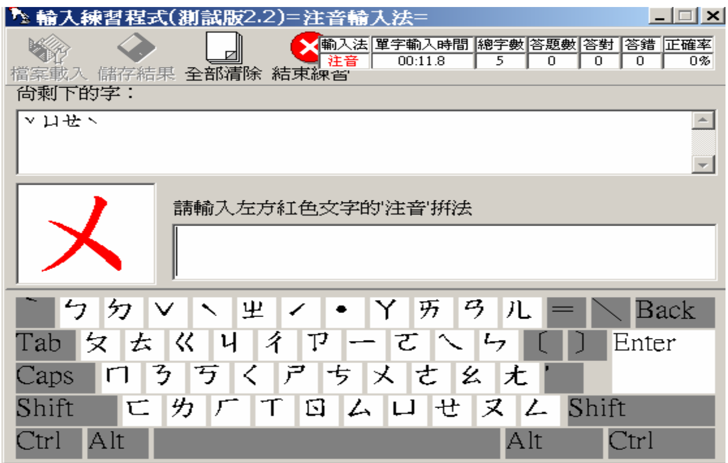


Fig. 1. The design of software screen Chinese Keyboarding Performance Evaluation

An on-screen Chinese Keyboard was designed on the lower half of the screen. A Chinese phonetic symbol will be generated randomly and shown on the left square of the screen. The participant was asked to type the indicated Chinese phonetic symbol on a standard keyboard or on-screen keyboard. The speed and accuracy of typing Chinese will be recorded and analyzed.

2.3 Experimental Design and Procedures

A group comparison experiment design was selected to examine the effectiveness of keyboard adaptations on increasing speed and accuracy for children with cerebral palsy. Speed and accuracy of typing Chinese phonetic symbol before and after keyboard adaptations were recorded and analyzed.

2.4 Data Analysis

The Wilcoxon rank-sum test, a non-parametric test was used to compare the speed and accuracy of keyboard inputs before and after keyboard adaptations. Statistical analyses were conducted using the Statistical Package for the Social Science, version 11.5.

3 Results

3.1 Keyboard Adaptation Strategies

Experimental results indicated that children with cerebral palsy often exhibit keyboarding difficulties which include insufficient ranges of motion of upper extremities to access across the whole standard keyboard, hitting the wrong key, repeating letters by mistakes, hitting the keys several times, and insufficient motor control to press two keys at the same time. The difficulties accounted by participants and the adaptation strategies suggested from the KAA software were summarized in Table 2.

Table 2. Keyboarding difficulties and adaptation strategies of participants

Keyboarding Difficulties	**Keyboard Adaptation Strategies**	**Participants**
Unable to reach across the whole keyboard	Mini keyboard	P1
Often hit the wrong key	Increase input acceptance time	P1, P3, P4, P5, P6, P7, P8, P9, P10
Letters repeated by mistake	Turn off key repeat Increase repeat time and delay	P1, P2, P3, P4, P5 P6, P9, P10, P12
Hit the keys several times	Increase post-acceptance delay	P1, P3, P4, P5, P6, P7, P9,P10, P12
Unable to hold two keys down at once	Sticky key	P1, P2, P3, P4, P5, P7, P8, P9, P10

3.2 Speed

The typing speeds before and after computer adaptations were indicated in Table 3. Children with cerebral palsy indeed improve significantly after keyboard adaptations ($Z=-3.061$, $p=.002$). For children with severe impairments, their improvements in typing speed increase much more than children with moderate impairments.

Table 3. Typing speed (inputs/minute) of the participants before and after keyboard adaptations

Participant	**Before adaptation***	**After adaptation***
P1	0.3	6.67
P2	3.7	11.67
P3	4.7	6.33
P4	9.3	11.67
P5	9.3	13.33
P6	11.3	15
P7	14.0	18.67
P8	15.0	17
P9	15.3	22
P10	20.7	23
P11	20.7	23
P12	23.0	26

Note: * means the average numbers of phonetic symbols the participant types per minute

3.3 Accuracy

Table 4 shows the typing accuracy of children with cerebral palsy before and after keyboard adaptations. The results did not differ significantly before and after adaptation for all the participants (Z=-1.947, p=.051). For most of the participants, their accuracy rate were pretty high before adaptation, therefore, they did not improve apparently after keyboard adaptation .However, for participant 1, 2, and 3, their

Table 4. Typing accuracy (%) of the participants before and after keyboard adaptations

Participants	Before adaptation (%)	After adaptation (%)
P1	7.7	95.2
P2	44	92.1
P3	17.6	95
P4	100	100
P5	100	100
P6	94.4	100
P7	97.7	94.9
P8	95.7	98.1
P9	95.7	98.5
P10	100	100
P11	100	100
P12	100	100

improvements of typing accuracy were profoundly which may due to their poor performance before adaptations.

4 Discussion and Conclusions

The purpose of this study was to compare the typing performances of children with cerebral palsy before and after keyboard adaptations. Typing difficulties and keyboard adaptation strategies of children with cerebral palsy were explored in this study. Speed and accuracy of typing were investigated systematically in this study. Typing speeds were increased significantly after keyboard adaptation in this study. However, accuracy rates of typing were not significant due to the ceiling effect of some participants. The accuracy did improve profoundly for some of the participants. The result of this study indicate that KAA is a valid reference tool when health professions working with children with cerebral palsy.

4.1 Limitations of Study

The results of this study provide preliminary supports on utilization of keyboard adaptation strategies for children with cerebral palsy. However, these results cannot be generalized to other children with cerebral palsy due to a limited number of subjects. Only twelve children with cerebral palsy participated in this study, their performance on computer adaptation cannot represent the variability of cognitive and motor abilities of all children with cerebral palsy.

In addition, group comparison may not an excellent way to study the typing performance of children with cerebral palsy since their performance were varied individually. This study is also limited to only pre and post adaptation performance was recorded per child. An intervention phase after keyboard adaptation may lead to greater improvements for those children with cerebral palsy.

4.2 Further Studies

A further research that validates the effective computer adaptation strategies in improving of life for persons with cerebral palsy is needed. Replication of this study using a larger sample of children with cerebral palsy, including children with different diagnosis and various ages should be developed to provide more information for clinical therapists working in this area. Single subject research design which provides sufficient periods for observing individuals with physical disabilities learning operating keyboards may be an alternative way to study this issue.

Acknowledgements

The authors gratefully acknowledge support from Nation Science Council, Republic of China, under Grants NSC 93-2524-S-182-003. The authors thank the study participants, their parents and occupational therapists for their cooperation. The authors also wish to thank Chi-Nung Chu for designing the computer software programs.

References

1. McCormack, D.: The effect of keyguard use and pelvic positioning on typing speed and accuracy in a boy with cerebral palsy. American J. of Occupational Therapy, Vol. 44 (1990) 313-315
2. Battenberg, J. K., Merbler, J. B.: Touch screen versus keyboard: A comparison of task performance of young children. .J. of Special Education Technology, Vol. 5(1) (1989) 24-28
3. Lau, C., & O'Leary, S.: Comparison of computer interface devices for persons with severe physical disabilities. American J. of Occupational Therapy, Vol. 47 (1993) 1022-1029

Light Spot Operated Mouse Emulator for Cervical Spinal-Cord Injured PC Users

Kazuyuki Itoh

Research Institute, National Rehabilitation Center for Persons with Disabilities, Japan
ito@rehab.go.jp
http://www.rehab.go.jp/ri/index.html

Abstract. The purpose of this study is to develop a mouse emulator system for cervical spinal-cord injured PC user. In this system, a laser illuminated point on a liquid crystal display is detected with image processing software developed for this system and the mouse cursor moves the detected point. The evaluation results indicate that the proposed method is comfortable for cervical spinal-cord injured PC user to operate GUI windows system.

1 Introduction

Persons with physical disabilities such as cervical spinal-cord injured (SCI) require assistive devices to use personal computers (PC), because their physical disabilities hinder them from utilizing a normal keyboard and a mouse. Although a mouse or head stick is the simplest assistive tool for them, the frequent use of these tools might cause additional disabilities or disease in other parts of their bodies. In order to prevent such side effects, we had developed a light spot operated keyboard (LSOK) for SCI PC-users since 1989. When it was initially developed the operating system (OS) was MS-DOS, so the development concept was to cast it as a keyboard emulate device. Thereafter, operating systems shifted to Windows, and the use of a mouse increased dramatically in graphical user interface (GUI) situation. LSOK have been clinically evaluated by SCI person, and have been found to be effective as a keyboard emulate device [1]. However, moving the mouse cursor is performed by illuminating a laser light onto a sensor corresponding to each shift direction (8 directions that include up and down, left and right, and diagonally) so, depending on the direction, the direction of the head and the line-of sight do not always match. In order to improve the mouse interface in relation to the LSOK user, we developed a system (light-spot operated mouse emulator) that detects laser illuminated points on a liquid crystal display (LCD) using image processing software and moves the mouse cursor to those sites.

2 Development Background

2.1 Existing Mouse Emulators

Mouse pointing devices for SCI PC users have been reported and some products are commercially available. These products use a switch operation for each direction and a joystick type operation [2]. And other devices perform a mouse operation by linking

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 973 – 980, 2006.

head movement with mouse cursor movement [3], thereby effectively utilizing the residual function of the SCI person. When using these devices, gain (setting of amount of mouse cursor move corresponding to head movement) and sensitivity are set. Problems occur, however, if these settings are not appropriate. With gain, strict calibration on commercially available products is not required, but the user must move the head so that the mouse cursor reaches the target site, such as an icon and menu bar. If the head can only be moved slightly, the mouse cursor may not always move to the periphery of the screen.

2.2 Adaptation to LSOK Users

We examined introducing the commercially available product Tracker-One [3] /Track-IR [4] for LSOK users (SCI person) for the purpose of using a Windows PC mouse, but the distance between the PC screen (19 inch) and user was about 90 cm, so there were some cases in which the mouse cursor did not move to the periphery of the screen using the gain adjustment the instrument was equipped with. A mouse pointing function that utilized a laser pointer was also desired because users were experienced with keyboard operations using a LSOK.

3 Detection of Pointing Sites on the Liquid Crystal Display Using Image Processing

If the user illuminates the laser pointer where he/she wishes to move the mouse cursor and the mouse cursor moves there, that should suffice for user interfacing. The present system is an application in which images uploaded with a USB camera are software processed, the laser illuminated point on the LCD is detected, and the mouse cursor follows the detected point. The procedure is as follows:

3.1 Use of Polarizing Filter

A polarizing filter is incorporated into the LCD on a PC. When an image is taken through a filter that has been rotated 90 deg, the LCD only appears black (Fig 1: Left, rotated 0 deg → Right, rotated 90 deg), and the LCD frame is easily detected.

Fig. 1. CRT image with rotated 0 deg and 90 deg PL-filter

3.2 Use of USB Camera

The LCD is imaged using a USB camera that utilizes the above mentioned polarizing filter, and the uploaded image is processed using software. The system is simplified by not using hardware such as an image processing board. Because laser light is reflected onto the screen periphery, two USB cameras are set up whereby an image from each one is synthesized to produce one image (640 x 480 pixels).

3.3 Software Image Processing

The uploaded image is converted to the binary data using the threshold setting and the LC frame is detected. The LC frame inside the synthesized image is not at this time a perfect square, so it is made into a square by calculating the left upper and right lower coordinates using the procedure below. Fig. 2 shows schematically the left upper part of the binary screen data. Each cell shows one pixel. '1' is actually "0xFFh".

Cy[m+1]- Cy[m]	m	Cy[m]								
	0	8	1	1	**1**	1	1	1	1	1
0	1	8	1	1	**1**	1	1	1	1	1
3	**2**	**5**	**1**	**1**	**0**	**1**	**0**	**1**	**1**	**0**
2	3	3	1	1	**1**	0	0	0	0	0
1	4	2	1	1	**0**	0	0	0	0	0
0	5	2	1	1	**0**	0	0	0	0	0
1	6	3	1	1	**1**	0	0	0	0	0
	n		0	1	**2**	3	4	5	6	7
	Cx[n]		7	7	**4**	3	2	3	3	2
	Cx[n+1]-Cx[n]			0	**3**	1	1	1	0	1

Fig. 2. Calculation of frame upper left coordinates

Cx[n] is the total value of the pixels in which the pixels in each row (Y=0～480) are '1' up to x=0～640. Similarly, Cy[m] is the total value of the pixels in which the pixels in each column (x=0～640) are '1' for Y=0～480. Next, the biggest differences between Cx[n] and Cx[n+1], and Cy[m] and Cy[m+1] were calculated, and the points where 10 pixels were added (or subtracted) to each were defined as the frame upper left coordinates. In the example in the figure above, the intersection point is where n=2 on the left edge and m=2 on the upper edge, and the coordinates are x=12 and Y=12. The situation is the same for the right lower coordinates.

Next, large bright pixels ('1') on the LC frame are illuminated point and the barycentric position of these pixels is considered the illuminated point of the laser pointer. The threshold was adjusted depending on the usage situation.

3.4 Mouse Cursor Moving

Fig. 3 is a schematic drawing of when the application is operated. The thick line is the PC screen. The application (long broken line) shows a screen in which the liquid crystal (LC) has been imaged. The mutual position of the LC frame (dotted line) detected within the screen and the illuminated detected point within the frame (o within the dotted line) is calculated, and then this is reflected onto the mouse cursor position (arrow) on the LCD to move it. In Fig. 3 a laser illuminated point has been detected at the top right of the LC frame, whereby the mouse cursor on the PC screen moves to the upper right position.

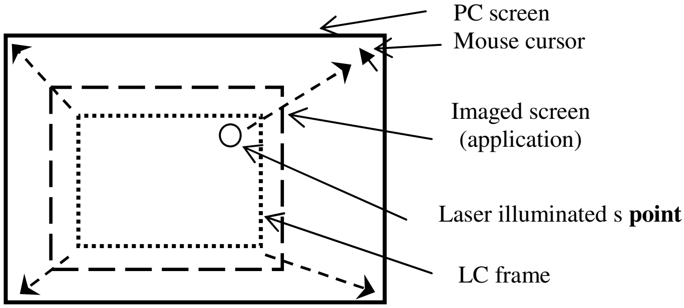


Fig. 3. Positional relationship between the calculated LC frame and laser illuminated point and mouse cursor position

Image processing occurs at about 10 FPS (frame/sec), and the coordinates for the laser illuminated point accumulate in the buffer at any time. For movement of the mouse cursor, the laser pointer movement amount m (dots/frame) for each frame is calculated, and that movement amount is performed after a continuous time (trajectory delay: T sec) has passed while the movement amount is no higher than the specified value M. By this process, if the laser pointer is moved quickly (M dots/frame or higher), then the mouse cursor will follow it after T sec after it has stopped, but if it is moved slowly, the mouse cursor will follow it if the speed is no more than M dots/frame). LSOK user perform mouse operations such as left and right clicking and dragging by illuminating the laser light on a special sensor set up at the periphery of the LCD, so to perform click operations by the above movement processing the mouse cursor does not follow movement even when the head is moved. It is possible to change the setting values for M (dots/frame) and T (sec) to meet the circumstances of the user.

3.5 Outline of System Set Up

The system is arranged as per Fig. 4. The sensors set up around the LCD are sensors for moving the mouse cursor used by the LSOK user. They are used to evaluate usage such as that discussed below.

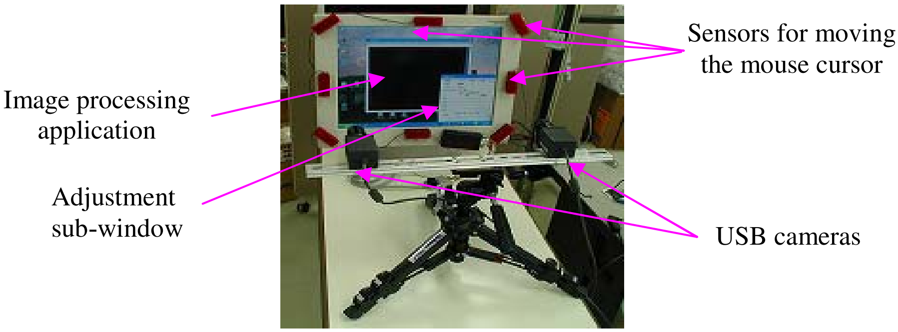


Fig. 4. System arrangement

4 Usage Evaluation

Usage was evaluated by comparing pointing operations that use the present system and pointing operations that use an optical sensor for movement. The parameters evaluated were the time it took for the mouse cursor to move to a target circle (T1 sec) and the time it took for clicking to be completed after the mouse cursor connected at first with the target (T2 sec).

As shown in Fig. 5, target circles appear in random order, one positioned in the center, and 30 positioned around it (in 8 directions, with three circles emanating out in each direction, and 6 circles around the periphery), so the subjects performed the operation by moving the cursor onto the target - which appeared in order from the center → periphery 1 → center → periphery 2 and so on - and selecting the target by hitting the laser on the optical sensor for left clicking. On the screen the selected target disappeared and only the target that should be next selected appeared. The movement time T1 was calculated as the time taken for the mouse cursor to reach the target after left clicking.

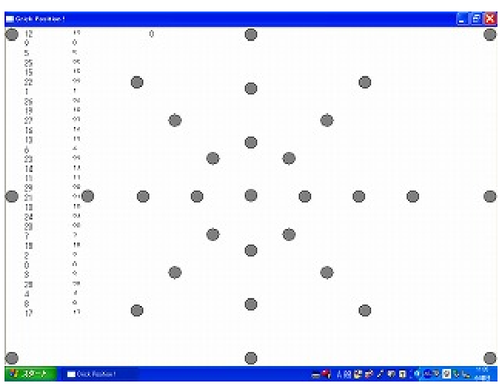


Fig. 5. Position of target circles

Subjects were one SCI person (SCI-A: 3 times by optical sensor), and three able body people(Able body -B: 1 person who used the LSOM system 1 time and the optical sensor 3 times; Able bodies -C & -D: 2 people who used the optical sensor 3 times). The SCI subject had used LSOK continuously for 13 years and was experienced at moving the mouse cursor using an optical sensor for mouse movement. The able bodied subjects had never before performed either type of operation.

Screen resolution was 1024 x 768 dots and the target circle size was 26 dots in diameter, which were the same values as one side of the 'Close/Minimize/Maximize' button in a standard display. The movement speed of the mouse cursor was set at 80 dots/sec. Sampling of mouse cursor positioning was performed every 50 ms. The values in the light-spot operated mouse system were set at 10 dots/frame for M and 0.4 sec for T.

5 Evaluation Results and Discussion

The performance of users was evaluated based on throughput (Amount of work processed in a specified time) values as defined below [5].

$$\text{Throughput} = \frac{\text{Log}_2(D/W+1)}{T1} . \quad (1)$$

D is the distance from the departure point to the target center, W is the target diameter, and T1 is the movement time. Units are bits/sec.

Table 1 shows the throughput values. The value in the table is the mean throughput value for the 30 target site selections. Calculations were made for each measurement until the target was reached (approach route) and from the target back to the center (return route).

When a movement was performed using the LSOM, the mouse cursor moved for a length of time that combined the time taken to lean the head towards the target and the tracking delay time, so a higher throughput value was obtained compared to when the optical sensor was used. Because on the return route the target position is always in the center and it does not need to be verified compared to the approach route, the throughput value for the return route is high. This is the same as when the optical sensor was used.

Table 1. Results of throughput (mean ± SD, bit/sec)

	Evaluate Times					
	1		2		3	
SCI-A	1.80±0.27	1.78±0.27	1.82±0.22	1.71±0.29	1.82±0.24	1.87±0.24
LSOM(B)	2.68±0.71	2.88±0.42	2.61±0.73	2.90±0.41	2.64±0.69	2.89±0.39
AB-B	1.10±1.42	1.09±0.38	1.16±0.44	1.16±0.35	1.21±0.37	1.10±0.33
AB-C	1.28±0.40	1.18±0.38	1.32±0.45	1.24±0.48	1.28±0.40	1.37±0.44
AB-D	1.04±0.39	0.88±0.29	1.22±0.41	1.09±0.30	1.20±0.31	1.14±0.41

The value was higher for the person with SCI than for the able body subjects because the former was experienced in movement operations using the optical sensor, and because he was good at fine control when the cursor approached the target. The same subject also rarely overshot the target.

Table 2 shows the mean time (T2) from when the cursor contacted the target to when the click was completed. When using the LSOM, the operation time was 30-60% shorter than the operation time for the optical sensor. As with throughput, the mean value of T2 was less for the subject with the SCI subject. However, this value also included recovery time when the target was overshot and the cursor traveled to the opposite side, so the result reflected the high level of operational experience of the SCI subject.

Table 2. Results of T2 (mean ± SD, sec)

	Evaluate Times					
	1		2		3	
SCI-A	1.2±0.6	1.3±0.8	1.0±0.5	1.0±0.5	1.0±0.6	0.9±0.3
LSOM(B)	0.9±0.4	1.1±0.4	1.2±0.3	0.9±0.3	0.9±0.6	1.0±0.3
AB-B	3.1±1.7	2.0±1.2	2.7±1.5	1.9±0.9	2.7±1.4	1.9±0.9
AB-C	1.8±0.9	1.1±0.8	1.5±0.9	1.7±0.8	1.4±1.1	1.0±0.7
AB-D	2.2±1.7	1.5±1.8	1.9±1.1	1.6±0.9	1.4±0.8	1.5±1.3

As a reference, Table 3 shows the total number of times the target was overshot in the movement experiment using the optical sensor. Compared to the record of a maximum of 36 times on the initial effort by the able-bodied subject B, the SCI subject-A only overshot the target 1 to 3 times, which indicates the level of experience of that person.

Table 3. Results of overshooting the target (times)

	Evaluate Times					
	1		2		3	
SCI-A	3	2	3	3	1	1
AB-B	36	20	28	14	28	5
AB-C	13	11	11	8	14	10
AB-D	13	11	16	15	8	14

As can be seen in Table 2 and Table 3, it could be expected that even when a person was inexperienced at operating the optical sensor, they could, with practice, acquire similar ability to the SCI subject who participated in the measurements. However, as shown in Table 1 and Table 2, big improvements in the throughput value and T2 (the time from mouse arrival at the target to clicking) were seen when moving the mouse cursor using the LSOM compared with the optical sensor. And, based on the evaluation experiment, operability was better when the LSOM was used.

6 Summary and Outstanding Issues

With the present system, because a laser pointer is illuminated onto a site where a mouse cursor is to be moved, gain adjustment is not required. The main problem is to combat noise such as the reflection of indoor light captured on the screen, so the setting position of the USB camera and an appropriate threshold setting are important. However, in a usage situation the demand is to reduce as much as possible the time and effort associated with settings, so it is necessary to further stabilize rough setting environments and provide a function for detecting point sites. The next step is to conduct a clinical evaluation among SCI subjects and investigate ways of improving the system.

References

1. Yasuo. S., Tadayuki S.: Improvement and Clinical Evaluation of Lightspot Operated Keyboard, Nat. Rehab. Res. Bulletin, no.10, (1989)115-119 (in Japanese)
2. http://www.kokoroweb.org/index_device.html (in Japanese)
3. http://www.madentec.com/
4. http://www.naturalpoint.com/trackir/
5. Edmund. F. L. and David. M. B.: Adaptive Software for Head-Operated Computer Controls, IEEE Trans. on Neural System and Rehabilitation Engineering, vol.12, no.1, (2004)102-111

Design and Implementation of a Chorded On-Screen Keyboard for People with Physical Impairments

Yun-Lung Lin[1], Ming-Chung Chen[2], Yao-Ming Yeh[1], Wen-Jeng Tzeng[1], and Chih-Ching Yeh[1]

[1] Department of Information and Computer Education, National Taiwan Normal University
162 Ho-Ping E. Rd., 1 Sec., Taipei, Taiwan
{harrison, ymyeh, memphis, ccyeh}@ice.ntnu.edu.tw
[2] Department of Special Education, National Chiayi University
300 University Road, Chiayi, Taiwan
mtchen@mail.ncyu.edu.tw

Abstract. The purposes of this study were to design an alternative on-screen keyboard for people with physical impairments and to evaluate the efficacy of the chorded input method. The approach of the on-screen keyboard is based on the human computer interface. It gives visual guide and instant feedback to show users where they can find the characters they need. The system has been designed with the principles of universal design. Three factors including the learning ability, efficiency of using and subjective satisfaction are considered as the usability evaluation. According to a preliminary study, the participant felt mastered the text input method quickly. An experimental evaluation on the typing performance of the subjects with muscular dystrophy will be measured under both scanning input mode and chorded input mode in the future.

1 Introduction

Computer technology has become an integral part in our daily life. It also helps people with physical impairments expand opportunities for educational success and improve their opportunities in both vocational settings and leisure activities. Text entry is widely regarded as one of the most frequent human-computer interactions. Using standard keyboard may seem pretty straightforward to most people. But for people with postural limitations, muscle weakness or limited movements, applying computer access are difficult. They encounter difficulties in various aspects when using standard keyboard, resulting in high error rates, fatigue and inefficiency. On-screen keyboards provide an image of a standard or modified keyboard on the computer screen by allowing users to select keys with a mouse, touch screen, trackball, joystick, or switch. Although the QWERTY layout is entrenched for standard keyboards and most of the on-screen keyboards are designed based on QWERTY layout. It is not a good user interface for people with physical impairments. The efficiency of the text entry depends on the suitability of control interface, which includes proper selection methods

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 981 – 988, 2006.

and keyboard layouts [1]. People who can manipulate only one hand, or people with more severe physical difficulties, may find the distance they have to travel to reach the whole keyboard restrictive. Such people may find that a more compact keyboard layout makes typing more efficient and less tiring for them. The purpose of this study was to design and implement an innovative on-screen keyboard for people with physical impairments which based on human computer interface. In the next sections the design of the on-screen keyboard and a pilot study will be discussed.

2 Design the On-Screen Keyboard

2.1 The Numeric-Based Input Method

Chorded keyboards use a combination of a few keys to create keystrokes for each letter. One clear problem is that typing text with the chorded keyboard is slower than using the QWERTY keyboard for novice user. The primary cost of the chorded keyboard is the extensive learning required associating the finger combinations with their corresponding actions, but after extensive practice, chorded keyboards have been found to support more rapid word transcription processing than the QWERTY keyboard, possibly due to reduced movement-time requirements [4]. Chorded keyboards are useful for people with limited arm range. Our chorded keyboard design efforts focus on input speed and learning time for people with physical impairments. It allows user to enter data and commands with one hand. We applied the numeric-based input method works by the user presses two numeric keys to generate character [3]. The interface design makes it simple and intuitive for users. To help users with their writing and editing, the letters will be pronounced when typed. It gives feedback to the user both visually through highlighting and through audio. As shown in Figure1, the first key press selects a group, and the second key select a letter, symbol or command in that group. For example, in order to type the letter ‘a’, the user first presses ‘4’ on the input device as shown to the left in Figure 1. The system will produce visual feedback to the user by highlighting the group of letters. And then, the user knows what group is activated. When the user presses ‘7’ (the second key), which selects the corresponding letter ‘a’ in the group. If a typing error was found, the user could use ‘0’ as a cancel button any time. For movement efficiency based on Fitts' law, the symbols and commands were configured by their frequency to reduce the fingers' traveling distance. For example, space is the most frequently used key and it is located on the center of the layout to reduce physical effort. The period mark is the most frequently punctuation, its corresponding two numeric keys are the same (2, 2). The travel distance of comma, question mark, and exclamation mark are shorter than other less frequently used punctuation mark. Although the frequency layout is theoretically physically fastest, for novice users, the entry speed is determined mostly by the needs to search and find target keys rather than by the amount of motor movement. The keyboard layout with alphabetical ordering will be faster initially for novice users.

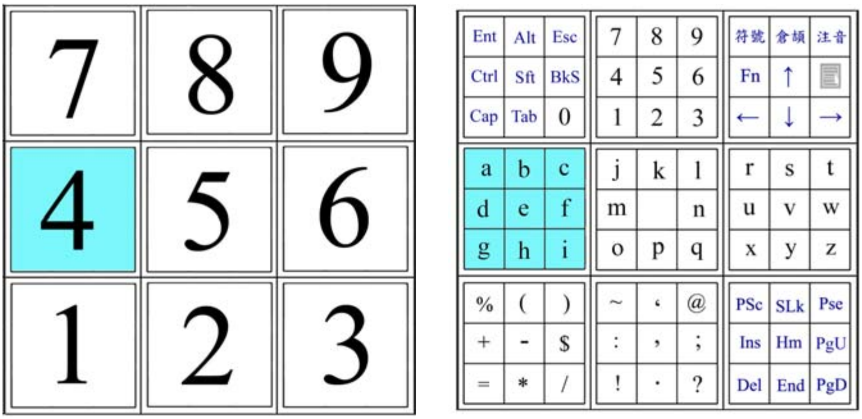
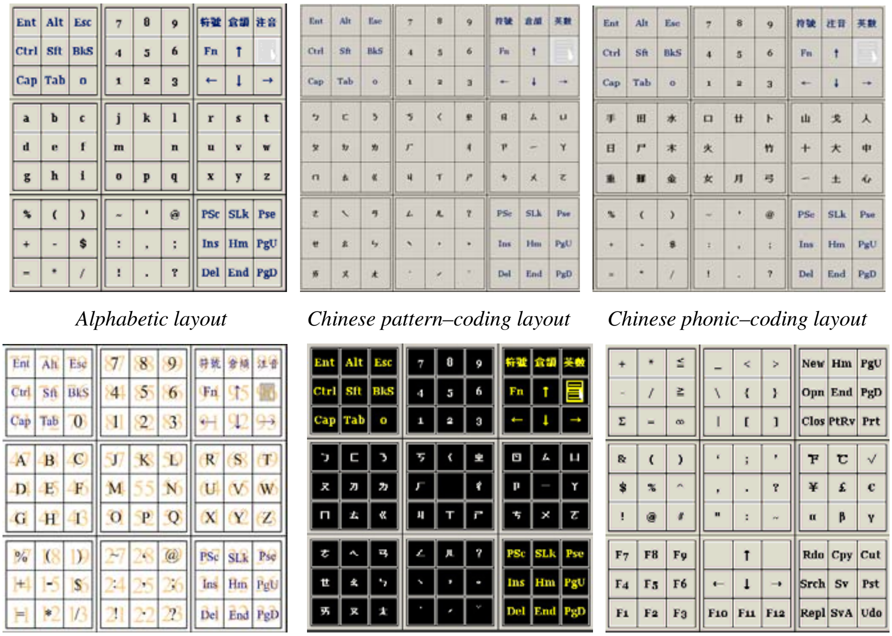


Fig. 1. Numeric-based Input method support visual feedback

2.2 Keyboard Layouts

Our chorded keyboard has been designed following the principle of universal design. It is displayed within a window that user can move, resize the layout and gives feedback to the user through audio. The system includes 9 layouts (international alphabetic layout, Scaffolding layout, Internet layout, 2 types of symbol layout, transparency layout, high contract layout and 2 types of Chinese Input layouts). The system can work with multiple access devices, such as keyboard, joystick, touch screen, touchpad, mouse, adapted switch. It also supports five alternative input methods (numeric-based mode, point-and-click mode, dwell mode, group-row-column scanning mode and joystick select mode). The users also could create their individual virtual keyboards by rearrange the special layout. Figure 2 outlines some keyboard layouts we implemented.

The scaffolding layout does not require much practice, and then can help novice user to reach a reasonable typing speed. The hint is superimposed directly on the target letter and the learner could pay attention to the same point. We suggest the novice user use the scaffolding layout at the first time. For visual impaired users, using the high contract layout, enlarging the size of the layout and using the option to pronounce letter may make a significant difference in the efficiency of their text entry task. Internet services also have particular benefits and potentialities for people with physical impairments. One of the useful applications for user is the World Wide Web (web). The keys on our internet layout are configured into three sets. One set for fast keys (1 in the Fig. 3), another for the command keys of browser (2 in the Fig. 3) and the last part is the same with the international alphabetic layout. Function key layout organized the common symbols, function keys and the common commands of the word processor (Microsoft word). The system includes a pop-up menu and allows reconfiguration as show in Figure 4.



Alphabetic layout *Chinese pattern–coding layout* *Chinese phonic–coding layout*

Scaffolding layout *High contrast layout* *function keys layout*

Fig. 2. The chorded keyboard layouts

Fig. 3. The internet layouts

Fig. 4. Pop-up menu

2.3 Group-Row-Column Scanning

Developing scanning mode assistive systems for people with physical impairments is not a new issue. The scanning approach is designed for people who fail to move the mouse pointer with any pointing device. Single switch automatic row-column scanning is the most common scanning approach for the people with severely physical impairments [5]. Beyond row-column scanning, we created group-row-column scanning based on our square layout. We predict that it is more efficient than row-column scanning with one switch activation. Figure 5 shows the display associated with the group-row-column variety of scanning. First, each group is automatically scanned in sequence. When the highlight arrives at the desired group, the user hits the switch (figure 5(a)). The highlight then advances scans each row in that group until the user hits the switch again (figure 5(b)). Finally, the highlight scans across each item in the selecting row until the user hits the switch (figure 5(c)). Scanning input method is less cognitive effort when compared to the chorded input method.

Fig. 5. Group-row-column scanning: the on-screen keyboard for single switch input

3 Usability Evaluation

To evaluate the usability, text entry tasks were performed to measure the typing performance of participant. Our initial trials for usability and efficiency have been restricted to one able-bodied college student. She was fluent in reading English and familiar with standard keyboard layout. The experiment started with an experimenter explaining the task and how to operate the chorded keyboard. And then, the subject took 15 minutes practice with the scaffolding layout and was asked to carry out typing test using a QWERTY keyboard and numeric-based input method. Each input method lasted 10 minutes. A picture of the test application is shown in Figure 6. An article is displayed on screen and the user is asked to type another copy and allowed to fix mistakes. Articles contain capital and lower-case letters, quotation marks and other punctuation. The test text for each trial was randomly selected from 42 articles, and both the typing speed and accuracy were recorded by the typing test program. Most people with physical impairments use only one hand for text entry, and novice user always use one finger for text entry task. In order to simulate the typing performance of single-digit user, the subject was asked to use a pencil instead of single finger. The subject must to carry out two keyboard patterns one trial a day, five days a week. We obtained ten measurements of the text entry speed as show in Figure 7. The performance of chorded-based, subject could type only 23.9 correct characters per minute in the first trial, and eventually up to 64.7 characters per minute in the last trial.

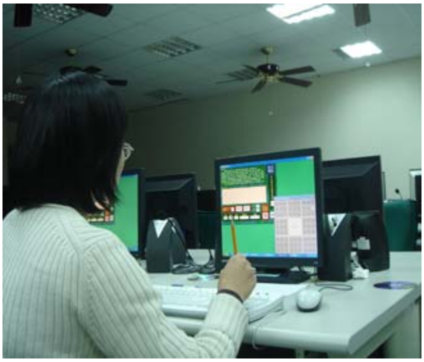


Fig. 6. Typing performance test

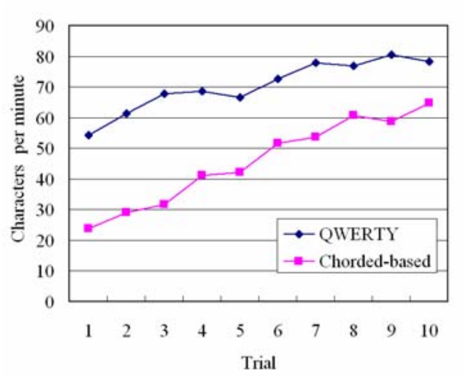


Fig. 7. Comparison of typing speed for QWERTY and chorded keyboard

Although the subjects' keystroke speed of chorded keyboard was slower than QWERTY keyboard, this might due to the subject's familiarity with QWERTY keyboard layout. However, the learning curve of chorded-based input method is very short. We also wanted to know how subject feel about the chorded input concept through post experiment interview. The interview was conducted after the final session. Subject was also encouraged to speak out and to write down comments or suggestions. The subject indicated that the numeric-based input method is easy to understand and the scaffolding layout is helpful for a novice learner. After using chorded keyboard for half of the third session, the subject did not even notice the concept any longer.

A potential limitation of this study is its reliance on a purely able-bodied subject, which may limit generalization the research results to the persons with physical impairments. However, we are confident that practice will give significant performance improvements and the chorded keyboard is more useful for some people with physical impairments.

4 Multiple Accesses

In order to make our chorded keyboard more useful for people with physical impairments, we have continually evolved the chorded input concept for user to multiple accesses. The system allows us to combine a wide variety of physical devices and language sets to form an interface ideally suited to the specific abilities of the user. The system can combine to combine other physical devices, such as N-fingers input device [2], USB calculator or PDA (Personal Digital Assistant) for some persons with muscular dystrophy (see Figure8 and Figure9). For persons with cerebral palsy, combine defined switches as input devices will help them computer access easily.

Fig. 8. USB calculator input device

Fig. 9. PDA input device

5 Conclusions

Although a large number of assistive technology products are available today, many people with physical impairments do not have adequate input devices. Most people with severely physical impairments have never been involved in computer usage throughout their life, interface designers must make the most of the metaphors involved. The need for participation in an emerging information society has let to several research efforts for designing accessibility solutions for people with physical impairments. Text entry for people with a motor impairment is a painfully tedious task. In this paper we present an alternative on-screen keyboard for them. We hope that the on-screen keyboard will contribute to creating accessible public computing facilities widely that can be enjoyed by people with physical impairments. Over the next few months, we will be improving the prototype and will then be offering it for people with cerebral palsy or muscular dystrophy.

We just begun our study of how the chorded keyboard can be used to help people with physical impairments, and there is much work to be pursued in many areas. We were able to put together several recommendations for the improvement of our on-screen keyboard based on the observations and the comments provided by the participant of the study. There are four implications and suggestions as following:

1. *Optimized keyboard layout.* The order of the letters in the scanning keyboards is crucial for their performance. To enhance the performance of our group-row-column scanning approach, the more frequently used keys could be rearranged together and to place the most frequently occurring characters in the most easily accessed locations.
2. *Word-prediction mechanisms.* Word-prediction is popular among devices that support text input. The on-screen keyboard embedded word-prediction mechanisms, characters prediction or abbreviation expansion will improve users' text entry speed.

3. *Embed intelligent system.* Reducing visual search for the selection will accelerate the typing speed. What's more, when an intelligent system that can recognize user's typing patterns is embedded, it will increase the performance.
4. *Voice recognition.* Voice recognition is an alternative input method. We will develop the voice command for our prototype in the future.

Acknowledgements

The first author expressed his gratitude to his mentor, Dr. Tein-Yu Li, who keeps inspiring him to finish this study. We also would like to thank Miss Ching-Hui Wang for her effort in coding this system.

References

1. Cook, A. M., & Hussey, S. M: Assistive technology: Principles and practice.(2nd Ed.), Baltimore: Mosby. (2002)
2. Ingmarsson, M., Dinka, D. & Zhai, S.: TNT-A numeric keyboard based text input method. Proceeding of the International Conference on Human factors in Computing Systems 2004, (2004) 639-646.
3. Lehikoinen, J., & Salminen, I.: An empirical and theoretical evaluation of BinScroll: a rapid selection technique for alphanumeric lists. Personal and Ubiquitous Computing, (2002) 141-150.
4. Richardson, R. M., Telson, R.U., Koch, C. G. & Chrysler, S. T.: Evaluations of conventional, serial, and chord keyboard options for mail encoding. Proceedings of the 31st Annual meeting of the Human Factors Society, Santa Monica, CA: Human Factors Society. (1987) 911-915
5. Zato, J.G.., Wangner, T., Barrasa, J., & Rodriguez, E.: Alternative web access toolkit for impaired users. International Journal of Rehabilitation Research, 28(1), (2005) 63-67

People with Motor and Mobility Impairment: Innovative Multimodal Interfaces to Wheelchairs

Introduction to the Special Thematic Session

Andreas Holzinger[1] and Alexander K. Nischelwitzer[2]

[1] Medical University of Graz
andreas.holzinger@meduni-graz.at
[2] University of Applied Sciences Joanneum
alexander.nischelwitzer@fh-joanneum.at

Abstract. Standard Interfaces have limited accessibility. Multimodal user interfaces combine various input and output modalities (including seeing/vision, hearing/audition, haptic/tactile, taste/gustation, smell/olfaction etc.), which are a classical research area in Human-Computer Interaction. One of the advantages of multiple modalities is increased flexibility in Usability. The weaknesses of one modality are offset by the strengths of another. For example, on a mobile device with a small visual interface and keypad, a word may be quite difficult to read/type, however very easy to say/listen. Such interfaces, in combination with mobile technologies, can have tremendous implications for accessibility and consequently, they are a potential benefit for people with a wide variety of impairments. Multimodal interfaces must be designed and developed exactly to fit the needs, requirements, abilities and different knowledge levels of the targeted end-users. It is also important to consider different contexts of use. However, in order to achieve advances in both research and development of such interfaces, it is essential to bring researchers and practitioners from Psychology and Computer Science together.

Keywords: Human-Computer Interaction & Usability Engineering (HCI&UE), Multimodal User Interfaces (MUI), Auditive User Interfaces (AUI).

Introducing Statement:
Today, together for better interfaces of tomorrow!

1 Accessibility of Computer Interfaces

Peoples' cognitive effort during the use of computer interfaces (especially of interfaces on mobile devices) can be dramatically high – especially for people with impaired abilities.

Complexity is increasingly becoming the barrier between advanced user interfaces and potential end-users [3]. In order to describe some effects of complexity and to form a possible basis for further research and practical implementation, the Cognitive Load Theory (CLT) is an ideal starting point [10]. Following this theory, one of our

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 989 – 991, 2006.

primary goals is to expand the abilities of the end-users working memory, thereby reducing their cognitive load.

According to Oviatt (2004) … *"there is currently no coherent theoretical framework that accounts for multimodal interaction patterns, which would be invaluable for proactively guiding the design of future multimodal interfaces to be optimally compatible with human capabilities and limitations"* [8].

However, not only is the cognitive ability a problem, many paraplegic people or people with motoric disabilities simply cannot hold the devices – such as mobile phones.

2 Multimodal Interfaces

Multimodal human-computer interface devices include joysticks, sip-and-puff-devices, touch screens and for output text-to-speech systems etc. Especially touch screen devices proved to be optimally suitable for disabled people [6,5].

These interfaces are interchangeable and can be adapted to the disability of a specific person. For example, a quadriplegic person can still control their tongue, therefore they are able to use sip-and-puff devices for example to control a Television set. Our premise is that end-users will not have to adapt to the interfaces, instead the interfaces must be adapted to the users.

"Performance improves further when adaptive processing tailors the interface to important user and environmental characteristics" as Oviatt (2004) emphasizes [9].

There are many user interfaces for severely disabled people, including [4,2,3]: Joysticks, Foot switches, Touch pads and Tongue pads, special sensors (such as bending sensors, pressure sensors), motion-detection and gestures recognition through sensors, eye gaze control, eye switch ([1,7,11]) and pneumatic switches (sip-and-puff).

However, to our experience most of these interfaces require costly hardware and a complex setup. Consequently, it is necessary to develop flexible, mobile and simple interfaces which can be used in everyday life and consequently require less complex and less expensive setup procedures.

Let us provide an example from our own experience:

A power wheelchair's input device can be used for controlling other functionalities over the wheelchair's steering control thus optimizing the human-wheelchair interface. Joystick control is widely spread among wheelchair users and is therefore already accepted as an input device (low barriers to use). The right interface simplifies many complex tasks for severely disabled people. This way they gain a new degree of personal freedom and quality of life. This augmentation of autonomy is accompanied by positive changes for all parties involved [12]. In this context, the main benefit for the handicapped users is that they can use this standard device for more tasks than just the control of the wheelchair.

3 Future Outlook

With reference to the enormous possibilities of multimodal interfaces and their potential benefits for end-users, a great deal of future research is necessary. The technology is already here, the know-how for optimum application is still lacking. Research questions include:

- How do the end-users really interact with multimodality in real-lfie, as against laboratory settings, and in what situational contexts are the end-users cognitively and/or motorically overtaxed?
- Which form of multimodality do the end-users prefer for communication?
- What communication patterns do they associate with tasks and dialogues?

However, this research must be carried out by both Psychologists and Computer Scientists in close cooperation. All research results, which are gained, must consequently be integrated into development, resulting in advantages at systemic level.

Let us bring together Psychology and Computer Science in order to proceed forward towards making our information society accessible to everybody, in order to make life easier for people who are less fortunate. It is bothour conviction and duty!

References

1. Blosser, S.R.: SCATIR Switch (`http://www.msu.edu/~artlang/SCATIR.html`), Michigan State University - Artificial Language Laboratory, 2005
2. Brisson, L.J.: Designing user interfaces for severely handicapped persons *Proceedings of the 2001 EC/NSF workshop on Universal accessibility of ubiquitous computing: providing for the elderly*, ACM Press, Alcácer do Sal, Portugal, 2001
3. Buxton, B.: Less is More (More or Less). in Denning, P. ed. *The Invisible Future: The seamless integration of technology in everyday life*, McGraw Hill, New York, 2001, 145-179
4. Hawley MS., Cudd PA., Wells JH., Wilson AJ and PL., J.: Wheelchair-mounted integrated control systems for multiply handicapped people. *Journal of Biomedical Engineering*, *14* (3)
5. Holzinger, A.: Finger Instead of Mouse: Touch Screens as a means of enhancing Universal Access. in Carbonell, N. and Stephanidis, C. eds. *Universal Access, Theoretical Perspectives, Practice, and Experience. Lecture Notes in Computer Science Vol. 2615*, Springer, Berlin, Heidelberg, New York, 2002, 387-397
6. Holzinger, A.: User-Centered Interface Design for disabled and elderly people: First experiences with designing a patient communication system (PACOSY). in Miesenberger, K., Klaus, J. and Zagler, W. eds. *Lecture Notes in Computer Science. Vol 2398*, Springer, Berlin et al., 2002, 34-41
7. LC-Technology. Eyegaze Eyetracking System (`http://www.eyegaze.com/`), 2005
8. Oviatt, S., Coulston, R. and Lunsford, R.: When do we interact multimodally? Cognitive load and multimodal communication patterns. in *6th international conference on Multimodal interfaces*, (2004), 129-136
9. Oviatt, S., Darrell, T. and Flickner, M.: Multimodal interfaces that flex, adapt, and persist. *Communications of the ACM*, *47* (1). 30-33
10. Sweller, J.: Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12* (2). 257-285
11. TashInc. Special Input Devices (`http://www.tashinc.com/`), 2005
12. Wobbrock, J.O., Myers, B.A., Aung, H.H. and LoPresti, E.F.: Text entry from power wheelchairs: edgewrite for joysticks and touchpads. in *6th international ACM SIGACCESS conference on Computers and accessibility*, (Atlanta (GA), 2004), ACM, 110-117

Accessible User Interface Framework for Severely Physically Disabled People

Michael Mahr and Alexander K. Nischelwitzer

FH JOANNEUM Gesellschaft mbH, research group Digital Media Technologies, degree programme Information Management, A-8020 Graz, Austria
{Michael.Mahr, Alexander.Nischelwitzer}@fh-joanneum.at

Abstract. This paper presents a software concept to allow severely physically disabled people to enter the information age. The software concept consists of a software framework which can be used to reduce the cost of adapting applications to special input devices. The software framework in particular allows different types of input devices for severely disabled people to be subsumed into one software interface. Thus if an application is extended to support this framework, the application automatically supports all current and future input devices for severely physically disabled people. It turns out that several software challenges need to be mastered in order for the framework to satisfy the tremendous need for adaptability of severely physically disabled people.

1 Introduction

1.1 Eliminating Barriers

Our world has undergone a major change from the industrial age to the information age. Anyone who does not have access to information is unable to take part in the information society and falls behind. This phenomenon is known as the 'digital divide' and is usually mentioned in the same breath with the Third World. However, even in Europe a group of people is endangered of becoming a victim of the digital divide: people with severe physical disabilities.

‘Everyone has the right to freedom of opinion and expression; this right includes freedom to hold opinions without interference and to seek, receive and impart information and ideas through any media and regardless of frontiers.’ [12]

Many disabled people are unable to exchange information, contrary to the above paragraph from the ‘Universal declaration of Human Rights’ of the UN. However, the barriers which deny disabled people this right are not geographical; the barriers are user interfaces designed for the average user.

This fact is especially frustrating, because the virtual world would allow severely physically disabled people to be normal in a certain way. The virtual world lacks physical representations and thus physical disabilities.

Unfortunately fulfilling all human rights for everyone is an impossible task. If fulfilling a certain right generates money or saves money, it will most likely be fulfilled sooner than other important rights. This is where this paper comes in: It presents a

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 992 – 998, 2006.

software concept for special input devices [13] in order to allow severely physically disabled people to enter to the information age.

1.2 The Target Group

Many applications are primarily developed for the average user and do not follow a user centered design [8,14] for disabled people. Having finished an application, developers sometimes deal with usability [10] and accessibility issues [9]. Most of the time, however, they do not consider this issue. That is why helper applications have evolved.

Helper applications translate the user interface into something a disabled person is able to use. Output-helper applications such as screen readers translate the output from the applications. Input-helper applications translate the input from special devices to simulate a keyboard and a mouse.

These helper-applications however are just a makeshift, since a lot of information and efficiency is lost during these translations. What is more, some applications, which do not adhere to standards, cannot be used with them at all.

A better solution would be to adapt the applications directly. However, modifying applications [2] that have been developed for the average user is a difficult and expensive task. That is why many companies choose to ignore the needs of disabled people.

Larger groups of people with the same disability have formed lobbies and have forced even large corporations to adapt their applications. Small groups, however, lack this lobbying power and remain unheard.

They consist of people with severe physical disabilities, who are unable to directly of indirectly control a mouse pointer and use special input devices instead, such as [11]:

- Sip and puff devices
- Blink sensors
- Joystick-like devices such as head-mice

If an application is adapted to all of these input-devices, the use-case for every device has to be considered separately and this is why implementation costs are very high.

The software framework presented in this paper allows subsuming all these devices into just one software interface [1]. Thus, if an application needs to be adapted, it just needs to be extended once to support all of the above mentioned input devices. The whole concept was developed for the 'MediaWheelie' [6] project, which gives disabled people more freedom and access (from environmental control [3,5] to information technologies).

2 Reducing Implementation Costs by Creating a Framework

2.1 Device Abstraction

The first step to reduce the implementation cost is to add an abstraction layer for input device types. This means that the application receives the user input from an abstract, unified device-interface instead of a real device. The application, for example, just needs to know how to communicate with a generic joystick-interface instead of implementing the protocol for several proprietary joystick devices. However, for the abstraction layer to work, every hardware-device needs a wrapper in order to be accessible through the generic device interface, also known as ‘driver’.

This variant allows applications to support new and future input devices [7] of the same type, without needing to be extended twice.

Abstraction layers for hardware devices are common practice for operating systems like Microsoft Windows [15], Linux or MacOS and are therefore not elucidated further.

Since many input devices are already supported by operating systems, it would be a waste of resources not to use these existing drivers for the framework and reimplement them for this framework. It is possible to use these operating system drivers by writing a universal driver which connects the device-stack of the operating system to one of the generic interfaces of the framework. This, for example, would allow all DirectX compatible joystick-devices to be used with this software framework, if a universal driver was written which connects to the DirectX joystick stack of the Windows operating system.

2.2 Data Type Abstraction

The second step in order to reduce the implementation cost is to add another abstraction layer which is positioned between the device abstraction layer and the application. The new layer abstracts the data which the application requires from the user. This means that the application just needs to specify which type of input it needs from the user as for example a piece of text, a number, or a direction. The rest is done by the framework.

This variant allows applications to support new and future input devices [7] of the same type and even future types of input devices, without needing to be extended twice.

However, in order for this to work the framework needs to convert the input into the requested data type. For example, if the user uses an USB joystick and the application requests the data type 'text' as shown in fig.1., the framework needs to convert the input from the generic joystick interface into the ‘text’ data type by using a converter plug-in.

A converter plug-in is similar to a driver. However, the converter plug-in reads from the device abstraction layer instead of a hardware-device. There are several algorithms for joystick-to-text conversion. Each of these would be implemented in a separate converter plug-in. If the application requires the user to input text, the configuration of the framework is parsed and the configured converter is loaded. From now on all joystick input from the user is converted into text until the user selects an escape command and thereby ends text-input. Then the text is sent to the application and the converter is unloaded.

This approach requires converter plug-ins from the user’s input-device type for every data type the application needs. A lot of converters seem to be required; however, the three data types 'Text', 'Number' and 'Direction' and thus three converter plug-ins are sufficient for most applications.

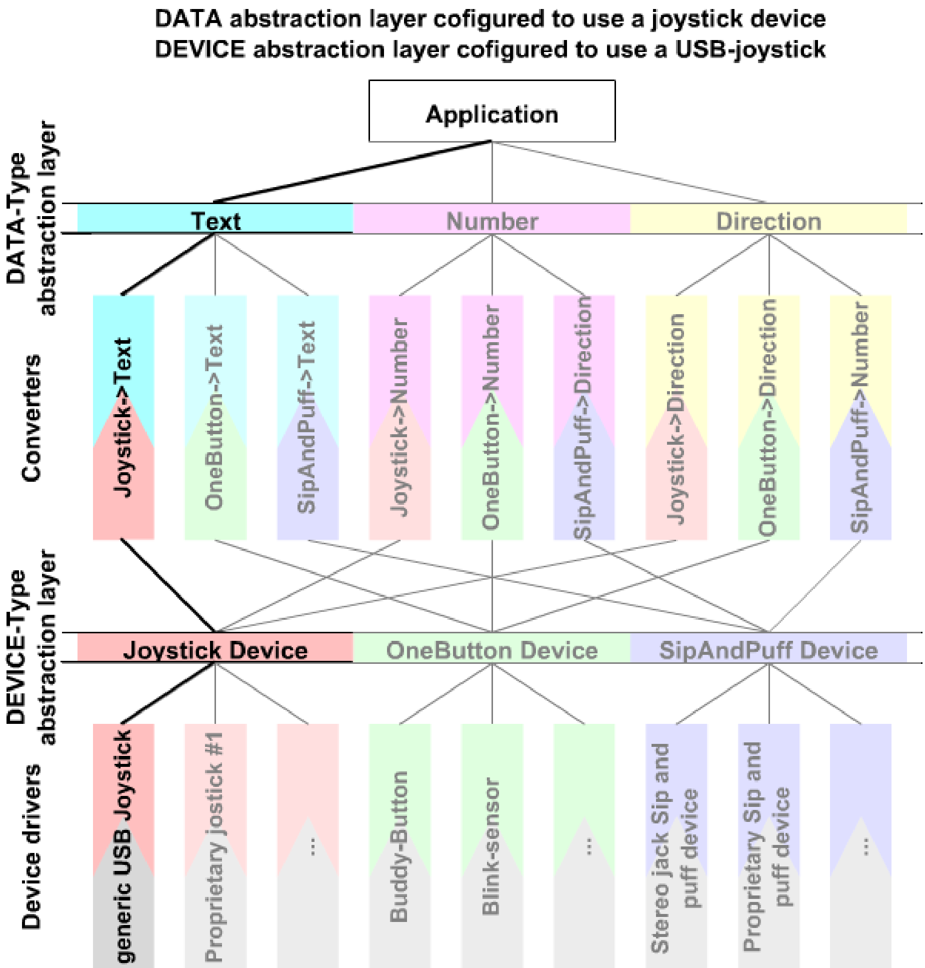


Fig. 1. Application requires the user to input text

Fig.2. outlines a Microsoft .NET implementation of the framework based on the same use-case as in Fig. 1.: a user writing text via a USB joystick.

There are four different types of plug-ins involved:

- Data types
- Converters
- Device types
- Device drivers

The data-type plug-ins define interfaces for the data abstraction layer. In this example the plug-in 'TextDatT.dll' defines the data type 'text' and allows the application to request text from the user.

Using plug-ins at this point is useful, because they allow application specific data types to be added later on. For example, if the application requires the user to select a button of a virtual remote control, it might take the user a long time to select a specific button when using a sip and puff device and the data types 'text' or 'direction'. However, if there was a data type called 'remote control button' it would be possible to write a converter which allows the user to select a button of the virtual remote control even faster. On the other hand this might also introduce dependency problems, if no appropriate converters exist in the framework for this data type. Therefore the applications should always provide a 'fallback' to the default data types 'text', 'number' and 'direction' in case no converter for the application specific data type can be found.

A central XML based [4] configuration file, in this case 'framework.xml', is used to define which converter plug-in is loaded if a specific data type is required by the user.

Converter plug-ins receive data from the device abstraction layer and send it to the data type abstraction layer. In our example this means that the converter reads its data from the abstract input device 'Joystick.DevT.dll', converts it to the data type 'text' and sends it to the application through the data type abstraction layer.

Each converter plug-in has at least one configuration file which serves two purposes:

- It defines specific attributes of the converter, such as its behavior.
- It defines which device driver is loaded when this converter starts.

The third kind of plug-ins are the device type plug-ins, which define the interfaces for the input devices. In this example the plug-in 'Joystick.DevT.dll' defines the interface for a generic joystick.

The last type of plug-ins are the device drivers, which implement a certain device type. The device drivers contain the code for actually interacting with the hardware device. They, too, have a configuration file. However, it is only used to configure the behavior of the driver.

When the framework starts, all device drivers should immediately be loaded and permanently kept in the memory until the framework shuts down. Since the converters are loaded on demand, the framework needs to iterate through all configured data types and to parse the configuration files of all associated converters.

3 Results

Implementing the framework shown in Fig.2. is a lot more difficult than the concept suggests.

If a disabled person for example is able to control two one button devices at the same time, they can work more than twice as fast as with just one one-button device. This, however, requires converter plug-ins to be able to read user-inputs from two input-devices at the same time.

It should also be possible for one physical device to consist of several virtual devices. For example a PC-joystick consists of the joystick switch which is represented by a virtual joystick device and a number of joystick buttons, which are each handled as a virtual one button device.

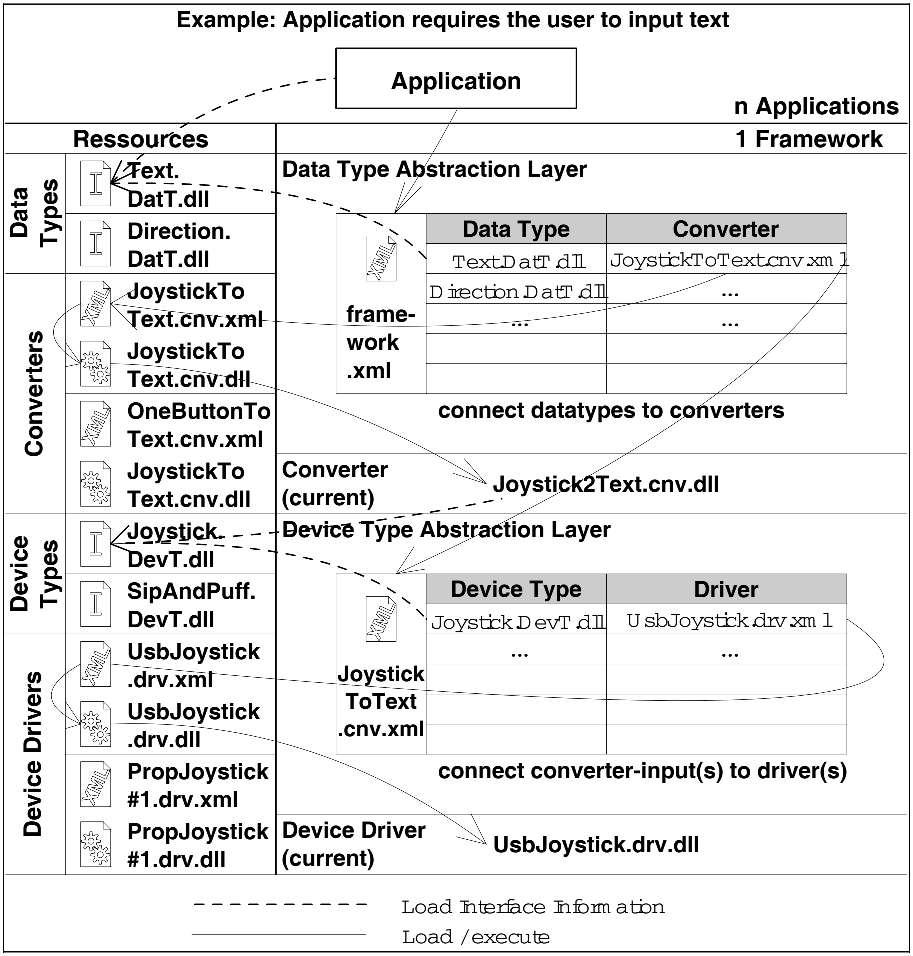


Fig. 2. Implementation of the framework

In future, it will be necessary to implement an output-framework which complements the input-framework. The output framework will be used to output different data types to the user via special output devices. However, it will also be used by the input-framework's converter plug-ins to give the users feedback on their interactions via well-specified interfaces.

Some input-devices also work as output-devices, as for example touch screens or game pads with actuators. This means that there is just one driver for both, the input- and output device, which implements the interface for the input framework and the output framework. As a consequence both frameworks need to work closely together and will probably be combined into one program.

Since the output framework will most probably become as complex as the input framework, it is likely that there will be eight different types of plug-ins. This results in a very complex individual configuration of the framework, where it will be easy to loose the overview, configure inconsistencies, and produce errors. The configuration files therefore need to be integrated with the structure of the plug-ins.

Ultimately the framework needs to be isolated from the plug-ins so that a bug in a plug-in can not crash the framework which as a result might crash the application too.

4 Future Prospects

In order for the framework to satisfy the tremendous need for adaptability of severely physically disabled people, several software challenges need to be mastered.

Ultimately the system will be easily comprehensible enough so that third-party developers are able to write drivers and converter plug-ins in order to adapt the system to match the needs of any disabled person.

References

1. Anderson, J., Fleek, F., Garrity, K. and Drake, F. Integrating Usability Techniques into Software Development. *IEEE SOFTWARE*, *18* (1). 46-53.
2. Brisson, L.J. Designing user interfaces for severely handicapped persons *Proceedings of the 2001 EC/NSF workshop on Universal accessibility of ubiquitous computing: providing for the elderly*, ACM Press, Alcácer do Sal, Portugal, 2001.
3. EIBA. European Installation Bus (`http://www.eiba.com/`), 2005.
4. Goldfarb , C.F. and Prescod, P. *The XML HandbookInc*. Prentice-Hall, Upper Saddle River (NJ), 1998.
5. Mace, R.L., Hardie, G.J. and Plaice, J.P. Accessible environments: Toward universal design. in Preiser, W., Vischer, J. and White, E. eds. *Design interventions: Towards a more human architecture*, Van Nostrand Reinhold, New York, 1991.
6. Nischelwitzer. MedienRolli Homepage (`http://www.medienrolli.com/`), 2005.
7. Nischelwitzer, A.K. Accessibility, the way to future interfaces and feedback possibilities - MediaWheelie as a best practice example *1st Usability Symposium; Empowering Software Quality: How can Usability Engineering reach these goals?*, OCG - Austrian Computer Society Vienna, Techgate, 2005.
8. Norman, D.A. & Draper, S. *User Centered System Design*. Erlbaum, Hillsdale (NY), 1986.
9. Raskin, J. We are all blind: Designing Interfaces for Accessibility. in *Lecture Notes of Computer Science, LNCS* 2004.
10. Shneiderman, B. Universal usability. *Communications of the ACM*, *43* (5). 84-91.
11. TashInc. Special Input Devices (`http://www.tashinc.com/`), 2005.
12. UNO Publications Board. Universal declaration of Human Rights, United Nations.
13. Vanderheiden, G.C. and Henry, S., Designing Flexible, Accessible Interfaces That Are More Usable by Everyone. in *CHI*, (Ft.Lauderdale (FL), 2003).
14. Vredenburg, K., Isensee, S. and Righi, C. *User Centered Design: an integrated approach.* Prentice Hall, Upper Saddle River (NJ), 2002.
15. Wigley, A., Wheelwright, S., Burbidge, R., Macleod, R. and Sutton, M. *Microsoft . NET Compact Framework. Core Reference.* Microsoft Press, Redmont (WA), 2003.

MediaWheelie – A Best Practice Example for Research in Multimodal User Interfaces (MUIs)

Alexander K. Nischelwitzer[1], Bernd Sproger[1], Michael Mahr[1], and Andreas Holzinger[2]

[1] University of Applied Sciences JOANNEUM
Alte Poststrasse 147, A-8020 Graz, Austria
{alexander.nischelwitzer, bernd.sproger, michael.mahr}@fh-joanneum.at
[2] Medical University of Graz
andreas.holzinger@meduni-graz.at

Abstract. For many motorically disabled people it is extremely difficult to operate keyboards. This is worst with tiny keyboards for example on mobile phones. This paper reports on research in multimodal user interfaces in order to assist disabled people to operate mobile phones or other equipment of daily life. Research in this area is gaining insight into problems and solutions which are generally applicable to assist the e-Society including elderly people or patients within a hospital. In the project described a so called MediaWheelie was developed, which is itself a common electric wheelchair which is extended by various multimedia devices. The focus was on usability and accessibility by following an end-user centered development.

Keywords: Multimodal User Interfaces (MUI), Accessibility, Audicons, Usability, Tacticons, Auditive User Interfaces (AUI), 2D audio interfaces

Introducing Statement:
The old interface is about what developers like;
the new interface is about what end-users really need.

1 Introduction

Motorically disabled people (caused for example by spastic tetraplegia, muscular dystrophy, cerebal palsy, Parkinson's disease, spinal cord injuries, etc.) are unable to use a conventional mobile phone. However, many people with motor impairments use wheelchairs and some of them use power wheelchairs [17]. We investigated multiple possibilities in interaction between human and wheelchair and consequently interfaces to connect them to the generic, modular interface of our so called MediaWheelie (`www.medienRolli.com`). One of the main goals of this project was to assemble standard hardware and software components in a modular way, in order to easily adapt the system to special end-user's needs; this was achieved by application of usability engineering methods [8].

Surveys showed that nearly 81% of power wheelchair users use a joystick as a control interface, followed by 9% using head or chin control, 6% using sip and puff, 4% using eye gaze [9], tongue pad, head, hand and foot switch controls. The very same

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 999 – 1005, 2006.

survey showed that 40% of power wheelchair users have difficulties controlling their wheelchair; a low percentage even finds it impossible to use it without assistance [6]. There are many user interfaces for severely disabled person, for instance [7,3,4]: Joysticks (4 and 8 directions; see Fig. 1.), Foot switches, Touch pads and Tongue pads, special sensors (like bending sensors, pressure sensors; see Fig. 3.), motion-detection and gestures recognition through sensors, eye gaze control, eye switch [1,9,16] and pneumatic switches (sip-and-puff; Fig. 3.). However, most of these interfaces require costly hardware and a complex setup. We tried to find a flexible, mobile and simple interface which can be used in everyday life and therefore requires a less complex and less expensive setup.

Fig. 1. MediaWheelie steering-joystick

A power wheelchair's input device can be used for controlling other functionalities over the wheelchair's steering control thus eliminating optimizing the human-wheelchair interface. Joystick control is widely spread beneath wheelchair users and therefore accepted as an input device (low barriers to use it). The right interface simplifies many complex tasks for severely disabled persons. That way they gain a new degree of personal freedom and quality of life is augmented. In a clinical context this augmentation of autonomy is accompanied by positive changes regarding physical and mental health of all parties involved [18]. One main goal for the handicapped user is, that he can use his "normal device" for more tasks than only control the wheelchair.

2 New Multimodal Interfaces

Standard feedback channels (visual, acoustic, haptic) are well researched, consequently we investigated the possibility of implementing earcons, icons and tactons [2]. We developed acoustic and haptic feedback channels, where the latter is based on skin stimulation [15] which are usable by power wheelchair users and evaluate their suitability. Many paraplegic people or people with motoric disabilities cannot hold devices, such as mobile phones steadily in their hands. However, with the MediaWheelie they do not need that ability and still can use the devices. Basically the

MediaWheelie consist of an electric wheelchair (from Sunrise Medical), human interface devices and many auxiliary systems. The human interface devices can be a joystick, a sip-and-puff-device or a touch screen and for output a text-to-speech system or a TFT display. These devices are interchangeable and adapted to the handicap of a specific person. For example, a quadriplegic person can still control his tongue and therefore use the sip-and-puff device. The most important aspect of the MediaWheelie is that the user can control all of the auxiliary systems with his favorite human interface devices. So the user does not have to adapt to the interface, instead the interface will be adapted to the user.

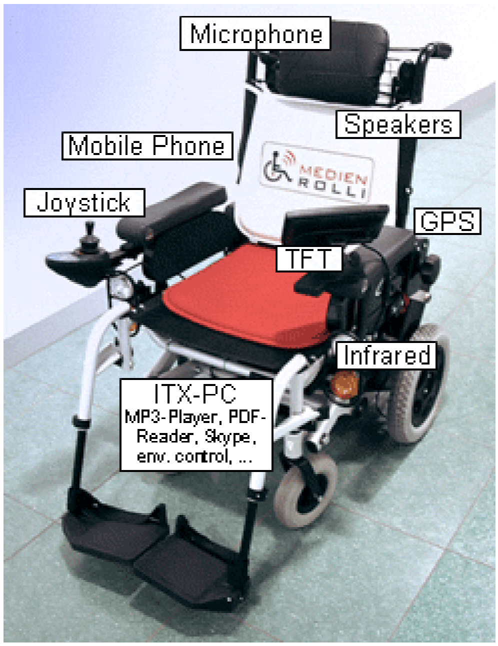


Fig. 2. MediaWheelie components include: Mobile phone (GSM, 3G), MP3 playback application, E-book (PDF) reader, GPS locator with SMS rescue system, Home device control (e.g. EIB [5]), Short message service (SMS), Internet telephony application (Skype), Infrared and Bluetooth communication, Wireless LAN (short area communication), Mobile UMTS internet connection (long area communication), DVD/Video Player (with optional device, touch screen)

The user of the MediaWheelie can make telephone calls, read E-books, E-mails and SMS, listen to mp3s, watch videos; additionally it allows the end-user to control all kinds of auxiliary devices including home automation systems (European installation bus by Konnex [5]). However, the most relevant and innovative aspect of the MediaWheelie is the integrated multimedia environment which allows adaptive usability and flexibility. This flexibility is possible through a complexity-reduced user interface, over which all multimedia devices and corresponding applications communicate. The modularity of the MediaWheelie allows all input devices to be interchangeable. So also tangible user interfaces [14] can be used, which allows new and intuitive input possibilities. The minimal output system used is an AUI (audio user interface) device. Special audicons [12] (auditory or natural audio icons) are used in combination with natural and synthetically generated speech.

3 Audio User Interface (AUI)

To start the MediaWheelie, the mobile phone and the miniature PC on the wheelchair need to be switched on. After around one minutes, the system is fully operable. By default the MediaWheelie is configured to use the joystick as input-device and text-to-speech for output. This setup is e.g. used for people who suffer from spastic tetraplegia, a motoric disability. In this case the TFT is not used and only displays the name of the current menu. The whole system is menu-based, which means, that all applications and auxiliary systems are listed in a cascaded menu. Whenever the user pushes the joystick, the MediaWheelie reads out the feedback of the action or the current location in the menu. Therefore it is impossible to get lost. If the user for example wants to switch on his TV-set, he has to go through the following procedure:

Table 1. Menu selection example to turn on a TV set

#	Joystick Input	Speech Output		Environment Action
		Feedback	Menu Info	
1	Down		Mobile Phone	
2	Down		Entertainment	
3	Right	Entertainment selected		
4			Media Player	
5	Down		Television	
5a	Down		Video Recorder	
5b	Up		Television	
6	Right	Television selected		
			On / Off	
7	Right	On / Off selected		
				TV turns on

The whole menu is defined in an XML structure and can be easily changed. Also the menu directions (left, right, etc.) can be changed with only one parameter. When a normal or a touch screen is connected to the system, the feedback will also be given in visual form to support a multimodal user feedback.

In the example before, the interface was a wheelchair joystick with 4 directions (4-way input). The modular software system can also handle other devices like special

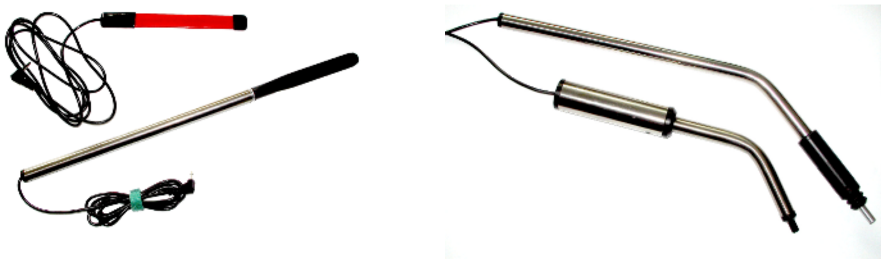


Fig. 3. Press Device (1 way) and pneumatic switches (2 ways sip-and-puff device)

buttons, 2-way devices like a pneumatic switch (Fig. 3.) or the eye with the scatir switch input device.

For example the sip-and-puff device can be used with the MediaWheelie to control the whole Microsoft Media-center. Depending on the task, the user has to select the navigation code like shown in Table 2. Green (↑) means puff and red (↓) stands for sip. The user must wait a short time between different commands. The "command end" time can be chosen depending on the user input velocity und timing. So it is possible also for handicapped people, to control things like video recorder, radio, TV set, mp3 player, slideshows and much more over a simple and standard sip-and-puff device via his wheelchair or directly from his bed.

Table 2. Navigation codes for sip-and-puff device for MS Media Center [10]

MEDIEN ROLLI

Navigation for sip-and-puff device

MediaWheelie Controls for MS Media Center

Navigation	
↑	OK
↓	Back
↑↑	Up
↓↓	Down
↓↑	Left
↑↓	Right

TV	
↑↑↑	Channel +
↓↓↓	Channel -
↑↑↓	Volume +
↓↓↑	Volume -
↓↑↓	Mute
↑↓↑	Info

Numbers	
↓↓↓↓	0
↓↓↓↑	1
↓↓↑↓	2
↓↓↑↑	3
↓↑↓↓	4
↓↑↓↑	5
↓↑↑↓	6
↓↑↑↑	7
↑↓↓↓	8
↑↑↑↑	9

Shortcuts	
↓↑↑↑↓	MCE start
↓↑↑↑↑	Power On/Off
↑↓↓↑	TV recording
↑↓↑↓	Live TV
↑↓↑↓	EPG
↑↑↓↓	DVD Menu

Videorecorder	
↓↓↓↓↓	Play
↓↓↓↓↑	Stop
↓↓↓↑↓	Pause
↓↓↓↑↑	Record
↓↓↑↓↓	Forward
↓↓↑↓↑	Rewind
↓↓↑↑↓	Replay
↓↓↑↑↑	Skip

Videotext	
↑↑↑↑↑	Videotext
↓↑↑↓↓	Input
↓↑↑↓↑	Delete
↓↑↓↓↓	Blue
↓↑↓↓↑	Red
↓↑↓↑↓	Green
↓↑↓↑↑	Yellow

4 Conclusion

The first main phase was the development of the MediaWheelie software architecture and the evaluation and implementation of different input devices for flexible environmental control. Alexander Weninger is using the system now since September 2005 and enjoys his new possibilities and freedom. His feedback together with the MediaWheelie logging data will be used for the next update and user interface improvements. One of the next goals will be the implementation of improved and fast text input methods, because this will help the users also to get a better chance for a job.

It could be shown, that with the MediaWheelie system, handicapped people can easily control devices, which they normally can not use. With the flexible software design, also a wide selection of input devices can be used. The next step will be the evaluation and usability testing (velocity, accuracy, learn ability, etc.) of the different input devices and the research for more complex input methods, for example optimized text entry (edgewrite, quikwriting [13], nokia T9, etc.) with eye movement or the wheelchair joystick.

Acknowledgements

The whole MediaWheelie project is supported by [11]:

- Government of Styria – Zukunftsfond Styria
- Sunrise Medical (Wheelchair support)
- AUVA - Allgemeine Unfallversicherungsanstalt
- ABB, EIB and IR Devices
- IBM, Speech Recognition, ViaVoice
- Telekom Austria, mobile Internet and Airtime

More information about the whole project can be found on the MediaWheelie Homepage: `http://www.MediaWheelie.com/` and `http://www.MedienRolli.com/`.

References

1. Blosser, S.R. SCATIR Switch (`http://www.msu.edu/~artlang/SCATIR.html`), Michigan State University - Artificial Language Laboratory, 2005.
2. Brewster, S. and Brown, L.M. Tactons: structured tactile messages for non-visual information display *Proceedings of the fifth conference on Australasian user interface - Volume 28*, Australian Computer Society, Inc., Dunedin, New Zealand, 2004.
3. Brisson, L.J. Designing user interfaces for severely handicapped persons *Proceedings of the 2001 EC/NSF workshop on Universal accessibility of ubiquitous computing: providing for the elderly*, ACM Press, Alcácer do Sal, Portugal, 2001.
4. Buxton, W., Foulds, R., Rosen, M., Scadden, L. and Shein, F. Human interface design and the handicapped user *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM Press, Boston, Massachusetts, United States, 1986.

5. EIBA. European Installation Bus (`http://www.eiba.com/`), 2005.
6. Fehr, L., Langbein, W.E. and Skaar, S.B. Adequacy of power wheelchair control interfaces for persons with severe disabilities: A clinical survey. *Journal of Rehabilitation Research and Development*, *37* (3). 353 - 360.
7. Hawley MS, Cudd PA, Wells JH, Wilson AJ and PL., J. Wheelchair-mounted integrated control systems for multiply handicapped people. *Journal of Biomedical Engineering*, *14* (3).
8. Holzinger, A. Usability Engineering for Software Developers. *Communications of the ACM*, *48* (1). 71-74.
9. LC-Technology. Eyegaze Eyetracking System (`http://www.eyegaze.com/`), 2005.
10. Lienhart Dietmar and Maitz Christian. MediaWheelie Input Device sip-and-puff. Nischelwitzer ed. *UPM - creative company*, Digital Media Technologies, Graz, 2005.
11. Nischelwitzer. MedienRolli Homepage (`http://www.medienrolli.com/`), 2005.
12. Nischelwitzer, A. and Janisch, G. AUI WALK – Online Usability Test Application for Performance Analysis of Auditory Mappings *HCI International*, Lawrence Erlbaum Associates, Las Vegas, USA, 2005.
13. Perlin, K. Quikwriting: Continuous Stylus-based Text Entry *Symposium on User Interface Software and Technology*, Proceedings of the 11th annual ACM symposium on User interface software and technolog, San Francisco, California, United States, 1998.
14. Ploder, Matzer, Nischelwitzer and Sprung, Magnfying Glass versus Mouse: A Usability Evaluation Comparing Tangible with Graphical User Interfaces. in *Multimedia Applications in Education Conference - Mapec 2004*, (Graz, Austria, 2004), 146-154.
15. Tactaid. Skin Stimulators (`http://www.tactaid.com/skinstimulator.html`), Tactaid, 2005.
16. TashInc. Special Input Devices (`http://www.tashinc.com/`), 2005.
17. Wobbrock, J.O., Myers, B.A., Aung, H.H. and LoPresti, E.F., Text entry from power wheelchairs: edgewrite for joysticks and touchpads. in *6th international ACM SIGACCESS conference on Computers and accessibility*, (Atlanta (GA), 2004), ACM, 110-117.
18. Wobbrock, O.J., Myers, A.B. and Kembel, A.J. EdgeWrite: a stylus-based text entry method designed for high accuracy and stability of motion *Proceedings of the 16th annual ACM symposium on User interface software and technology*, ACM Press, Vancouver, Canada, 2003.

NIBLUM: Non-Invasive Bluetooth Mouse for Wheelchair Users

Roberto Casas[1], Marcos Quilez[1], Borja Romero[2], and Oscar Casas[1]

[1] Instrumentation, Sensors and Interfaces Group,
Castelldefels School of Technology (EPSC),
Technical University of Catalonia (UPC), Spain
{rcasas, marcos, jocp}@eel.upc.edu
[2] B&J Adaptaciones, Barcelona, Spain

Abstract. Being autonomous is one of the biggest challenges for many people with disabilities. While wheelchairs enable individuals to move freely, computers allow them to communicate, work, etc. by their own. In many cases both devices are controlled by two different interfaces of the same kind -two joysticks- and people need the assistance of a second person to switch between them. In this article we describe how any person that drives a wheelchair using a joystick, can control the pointer of any computer using the same driving joystick and without being helped by anyone. This device, NIBLUM, uses Bluetooth to get a wireless and transparent connection with the computer. As it is not necessary to modify the wheelchair, its manufacturer's guaranty is not voided, something essential to enable NIBLUM's real usage.

1 Introduction

Computers are becoming more and more present in our lives; personal computers allow us to communicate, work, learn and perform many leisure activities. Portable devices such as laptops, tablet-PCs and PPCs, are also more habitual because they enable the user to perform all these actions almost everywhere. Advances in technology such as miniaturization, power consumption and weight reduction or wireless communication systems have outstandingly contributed to this.

Disabled people can take advantage of the spread use of computers thanks to adapted software and peripherals. Not only do computers help them with those habitual tasks, but they also allow them to do some other basic activities that otherwise would be impossible. People with speech dysfunction can make sentences that are synthesized and read by the computer, those suffering from cerebral palsy can control their environment through the computer, etc. In addition, portable devices let the users take advantage of all these features everywhere. As a result, computers give more independence to people and improve their quality of life.

The way people interact with computers is an important issue not universally solved. Mice, keyboards and pointers are the most usual input interfaces; lamentably not everybody is able to control them properly. Both, PPCs and PCs, often have operating systems based in graphical user interfaces such as Linux or Windows. Then, pointing devices (mice and pointers) become essential for a full control of software.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 1006 – 1013, 2006.

However, people with severe motor-impairments are unable to use those standard interfaces and adapted ones are needed for pointing and typing.

Numerous alternative ways of controlling the mouse can be found in the bibliography. Sporka et al. process sounds acquired through the computer's sound card to perform basic mouse pointer operations: click, double click and movement in the two axes. Main advantages of this system are low computation power needed, short learning curve, easy installation and no-necessity of special devices [1]. In this line, using speech (elaborated sound) recognition techniques such as IBM ViaVoice or Dragon NaturallySpeaking are also suitable, but mostly used for textual input [2].

Vision technology is another possibility widely studied. The analysis of video images taken from a camera to determine the user's gestures or head position is an alternative requiring low infrastructure and high computation resources [3]. To improve the performance while reducing computation load, alternative non-visible lighting sources (infrared) are used to create a controlled illumination scenario. Adjouadi et al. use a commercial system that uses this operation principle to track user's pupil and build an integrated, real-time computer interface that allows controlling most of the Windows applications [4]. Wearing specific equipment such as reflective targets or special glasses makes the image processing simpler. Prentke Romish's Tracker 2000 is a commercial system using this approach [5]. Evans et al. process the radiation emitted by a small device attached to the user's head to determine the head position [6].

Physiological sensors can be used to translate brain or muscle signals into cursor movements. Lamentably, it is hard to uniquely associate EEG or EMG patterns with mouse actions using these techniques [7].

Inertial sensing is another technology broadly explored; tilt sensors and gyroscopes sense inclination and rotation enabling coherent mouse operation. Chen uses inclinometers attached to the user's head to detect lateral and vertical movements [8]. Bending the head towards the shoulders to move the cursor left and right, is a movement that can become uncomfortable after some time [6]. To overcome this, detecting the rotation of the neck using a gyroscope is more appropriate. [9].

The most used and simple systems are joysticks with different interfaces and movement interpretation to adapt its handling to the abilities of the user [10].

Each system has its own advantages and drawbacks, but it makes no sense to discuss about which one is the best because it will heavily depend on the abilities of every user. Depending on the physical operation principle used the target population will vary. For a person that cannot talk and move autonomously, voice recognition and inertial systems will be useless, no matter the quality of these systems. However, an eye tracking system can maybe enable him/her full control of the mouse pointer.

2 Requirements of the System

Assistive technology (AT) devices are highly user-dependant as each disability hinders people in a different way. Thus, prior to any development, it is necessary to define the target audience that will use our system [11].

We are going to focus on people that are able to control a wheelchair using a joystick by themselves (no matter if they use their hand, chin, etc.) and that are not able to use standard interfaces to control intelligent devices such as PCs or PPCs. A broad range of people fits this description; those with severe motor-impairments (neural degenerative diseases, spinal-cord injuries, etc.) for instance.

We can show two different scenarios to illustrate the necessity we will satisfy with our system. The first one is a quadriplegic man that controls his wheelchair with a chin-joystick. He wants to use the PPC attached to the wheelchair when stopped, but he is not able to handle the pointer; thus he uses another adapted joystick (instead of the driving one) that has to be installed by someone else. The second user is a woman able to control the wheelchair with a hand-joystick, but that cannot move the arms to handle other objects by herself. She also has two computers, one at home and another one at work. As she is not able to control a standard mouse and keyboard, uses an adapted joystick, but someone has to move her arm so she can grab it. As both users are able to operate the wheelchair, the driving joystick could be used to operate the computer also. Unfortunately, wheelchair manufacturers do not offer this feature and manipulation of the joystick's hardware is not possible because it will void the guaranty. As a result both people lose their autonomy and independence because they need some help to get the computer or PPC interface prepared for operation. It is to say, they are not able to communicate, work, play, etc. unless someone else helps them with their interface devices.

The main objective of our system is to enable people that use a joystick (no matter the type) to operate their wheelchairs, to seamlessly control a computer or PPC mouse with the same device. This would avoid the need of someone else to change the driving joystick by a different interface.

The definition of the target audience, together with the definition of how the user will interact with the system [11], determine the requisites the system should fulfill to develop a functional and really useful device. The desired features are the following:

- The movement of the pointer should be controlled with the wheelchair-joystick with independence of its type and position. Guaranty of the wheelchair's manufacturer should remain valid, thus physical joystick manipulation is not acceptable.
- The handling of the wheelchair should not be affected by the device and if possible, the user should not need to wear additional equipment.
- Clicking of mouse buttons should be integrated in the system; nevertheless, other mouse actions such as double click or drag, should be configurable and allow adapted interfaces.
- The connection among the system and the computer should be transparent (minimum configuration required) and wireless.
- Because many types of computer devices are used by people with disabilities, it should be the most compatible possible, at least with desktop and tablet PCs, laptops, and PPCs.
- It should require the minimum custom software in the host computer.
- It should be reliable; remain unaffected by external interferences and automatically recover from failures.
- The power consumption should be comparable to the one of commercial computer peripherals. If not, battery maintenance should be low.
- It should be cost-effective.

3 System Design

To cover all the requisites detailed before, we have designed a cost-effective wireless controller for the pointer mouse that works in any computer with Bluetooth capabilities. This is done with the same driving-joystick of the wheelchair, without affecting its normal handling and with no manipulation of its hardware. We call it NIBLUM (Non-Invasive Bluetooth Mouse for wheelchairs).

Interface (Movement and Actions). The interface defines how the user will interact with the device. This is how he/she will move the pointer over the computer screen and how will perform the mouse operations (click, double click, drag, etc.). Regarding to the movement, it is a requisite that it is controlled with the same wheelchair's driving-joystick (not any other joystick); thus, there are no alternatives on this: The pointer will move following the joystick. Left and right displacements will move the pointer over the screen in the same direction. Moving the joystick to the front will cause the pointer to go up, while pulling it back will make the pointer go down.

The performance of actions is not tied to any functional requisite; each case will have its own necessities depending on the user abilities. Trevin and Pain make a wide analysis about the difficulty of performing habitual mouse actions by people with various motor disabilities [12]. Based on this study, we implement five actions that will be controlled separately: right, left and middle click, double left click and pointer drag & drop. These actions will be performed using switches; to allow user adaptation we found the best solution to have, besides the habitual three buttons, jack connectors for five inputs so that different actuators can be connected. This open architecture will take advantage of the user abilities to perform the pointer actions. For example, if the user can use the knees, chin, hands, etc. the system could have one to five buttons to perform the desired actions.

There is an additional switch -CONNNECT- that does not perform any standard mouse action, but it is necessary in this case. It is used to switch the destiny of the joystick's movement from the wheelchair to the computer and vice versa. It will cause both connection/disconnection of the wheelchair's motors and wireless connection/disconnection of NIBLUM to the computer.

Sensing (Processing and Calibration). As the pointer movement will be governed with the wheelchair's joystick -that indeed should not be manipulated- we sense its inclination attaching to it a tiny 2-axis accelerometer. The accelerometer is placed with the two sensing directions perpendicular to the joystick's main axis. As accelerometers measure acceleration, if static and parallel to the earth's surface, it will detect zero acceleration. When operating the joystick, as the sensor is attached to it, the gravity component will appear in the sensing axes and inclination can be measured.

User's movement will not interfere sensor's measurements because he/she will always be static when using the wheelchair's joystick as a mouse; the wheelchair's motors have been disconnected when pressing the CONNECT button. However, the accelerometer will not be always parallel to the ground in the joystick's resting position; the user can be stationed in a slope or the driving-joystick not be in the armrest (mouth or chin joysticks). To overcome this, we calibrate the neutral position and calculate the relative inclination from this position. When totally horizontal, the calibration point will coincide with the zero inclination point. In other cases, we obtain an offset inclination in both axes. This offset is subtracted to the measurements in normal mode to obtain the joystick's deviation from is resting position.

The resting position has to be calibrated every time NIBLUM starts to be used. The user can be in a different position (slope) than the previous time it was used or the joystick can have changed its position. As this calibration should not need the action of a second person, we perform it when the user presses the CONNECT button. Thus, at this moment we disconnect the wheelchair motors, establish the wireless link and calibrate the neutral position.

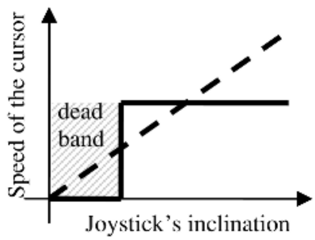


Fig. 1. Relationship among the speed of the cursor and the joystick's inclination in basic (solid) and advanced (dotted) modes

Once we know the joystick's relative inclination from its resting position, it is necessary to calculate the corresponding pointer displacement over the computer screen. Two operational pointing modes are described in the bibliography: relative and absolute [6]. Absolute mode relates the position of the pointer with the joystick; for example, in the resting position it will be in the center of the screen. This mode is more suitable for head pointing devices and for people with good movement control. When using joysticks, relative pointing is more appropriated; we also do. It works moving the cursor while the joystick is displaced in the desired direction; once it comes to the neutral position the pointer is stopped. The displacement speed can be constant or proportional to the joystick's inclination. First case provides easier handling and more robustness when the user has difficulties controlling the joystick; for example, caused by tremble. On the other hand, it is slow and less efficient for people without joystick's interaction problems and with some experience; here, speed proportional displacement is best suited.

Tremor and other involuntary movements can provoke moving the cursor when not desired and difficulty to point accurately. In these cases, besides using constant speed displacement, we define a dead band around the joystick's resting position. To start moving, the user must move the joystick beyond this band.

We define two control modes: basic and advanced. The solid line in the figure 1 represents basic mode with constant speed movement of the cursor and a dead band around the resting position. In advanced control mode, dotted line in figure 1, the cursor displacement speed is proportional to the joystick's inclination

Each person has different abilities, which are habitually improved through training or degenerated because of disease evolution. To broaden the possible user's population and extend the device's life to the maximum, it is possible to easily switch among basic and advanced modes with a switch.

Communication. The communication among NIBLUM and the host computer must be wireless because the user might not be able to plug any cable and mandatory assistance of a second person is not desired. We use Bluetooth because, as we will see, it allows easy fulfillment of many of the requisites in section 2.

Besides being wireless, it is a standard widely integrated in many computer devices. This enables immediate interoperation with most of the laptops and PPCs nowadays without adding any hardware. In other cases, giving a PC Bluetooth capabilities is easy with a cheap USB dongle.

Regarding to the link setup, it is desirable that both devices get automatically connected when the user presses the CONNECT button. Thus, any configuration of the software each time the user needs the computer should not be needed because it will reduce user independence. As Bluetooth was designed to replace wired point to point, short distance connections (serial or USB), issues such as discovering, authentication or connection are taken into account and solved in the standard [13]. Only the first time NIBLUM is used with a computer device, it is necessary the assistance of a second person. He/she must configure the software and pair both devices to allow further automatic link establishment. Once paired, the computer and NIBLUM will quick and automatically recognize each other and get connected. Disconnection does not need any preparation; it can be done powering off the device or moving out of range.

Software. Software is an important issue because NIBLUM is intended to work with a wide variety of computer devices. This implies developing software for different platforms and overcome compatibility problems. Fortunately, all these matters are avoided using Bluetooth. Bluetooth is not only a wireless communication protocol; it also defines operation profiles that implement concrete applications: serial port, headset, hands-free, etc. [13]. This guarantees interoperation among all the devices implementing the same profile no matter the manufacturer. We will use Human Interface Device (HID) profile. It defines the protocols, procedures and features that shall be used by human input/output devices (mouse, keyboard, joystick, etc.) communicating with a host (computer device). Thanks to that, we will not need to develop any specific software for the host side because most of systems with Bluetooth capabilities include HID profile.

4 Implementation

The device has been designed to be highly configurable to best fit the needs of every user. The sensing part (accelerometer) is attached to the wheelchair's joystick in order to sense its inclination. This sensor is plugged to the main unit that acquires its information, performs calibration and calculates the pointer displacement depending on the selected operation mode (basic or advanced). It also includes the operation mode switch and the CONNECT and action buttons. The operation of all the buttons can be overridden by using alternative adapted actuators connected to the existing jacks. All these operations are carried out by a microcontroller that sends the mouse operations to a Bluetooth module that implements the HID mouse profile. In figure 2 we show the block diagram of the system.

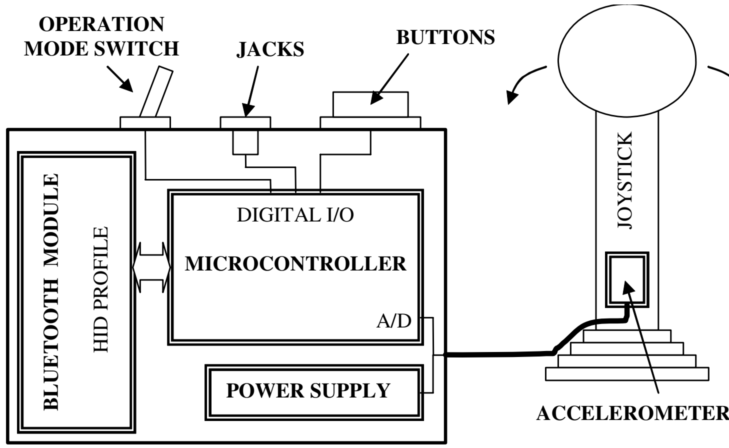


Fig. 2. Block diagram of the system

As a portable device, power consumption is a key issue that may turn the system useless. In our case, the current consumed by the accelerometer and the microcontroller is around 1 mA in active mode. However, the Bluetooth module uses up to 50 mA when transmitting and around 30 mA when connected in active mode even without data exchange. To reduce this energy devouring, we use Bluetooth sniff mode [13]. This way, the device only participates in the connection (listen and send data) for a short instant at regular intervals; the longer the intervals are, the less average power will be used. As a result, regarding to power consumption, our device is in the range of commercial mice. Percept Technology Labs made a study comparing Microsoft and Logitech wireless mice resulting in an 11 mA to 24 mA in continuous operation - we consume around 15 mA- [14]. In idle mode we go to 4 mA while others consume from 1 mA to 1.3 mA. When not connected all the devices fall below 300 uA. These results indicate that our device is suitable to be used as a computer mouse in terms of battery life time.

5 Discussion

The device has been tested with 20 people, of whom five were people with disabilities. All of them successfully used the device. The most appreciated feature was the possibility of switching the joystick function from handling the wheelchair to controlling the pointer without the assistance of a second person. Dedicated buttons to perform double click and, specially, drag&drop actions were also highly valued. The duplicity in the controlling modes showed its usefulness; not only four users chose the basic control mode as the best suited, and one of them was unable to operate the device in the advanced mode. On the contrary, the remaining users graded the basic mode as slow and in some cases even exasperating.

The device developed fulfils all the requirements stated in section 2. It detects the joystick's inclination no matter the position of the wheelchair and produces the corresponding movement of the pointer on the computer screen. To complete a real mouse performance, actions like right/middle/left click, double click and drag&drop are carried out with adapted interfaces that can be connected to NIBLUM. All this is done without modifying the wheelchair and with no limitation in the computer type to be used (PC, laptop, PDA), it only has to have Bluetooth capability. It can also work with any operating system because no custom software has to be used; only the standard Bluetooth HID profile needs to be installed. Regarding with practical aspects, it has a competitive power consumption that makes it suitable to be fed with normal batteries. Cost is another key issue that determines if it can be used in real situations. NIBLUM is going to be commercially available and compared to similar products it will be much cheaper.

References

1. Sporka, A.J., Kurniawan, S.H., Slavík, P.: Acoustic Control of Mouse Pointer. Universal Access in Information Society, Springer-Verlag Berlin Heidelberg New York (2005) 1-9
2. Basson, S.: Speech Recognition and Accessible Education. Speech Technology Magazine 7(4) [On-Line], 2002. Available at: http://www.speechtechmag.com/issues/7_4/avios/.
3. Tu, J., Huang, T., Tao, H.: Face as Mouse Through Visual Face Tracking, Proceedings of the Second Canadian Conference on Computer and Robot Vision (2005) 339-346
4. Adjouadi, M., Sesin, A., Ayala, M., Cabrerizo, M.: Remote Eye Gaze Tracking System as a Computer Interface for Persons with Severe Motor Disability. Lecture Notes in Computer Science, Vol. 3118. Springer-Verlag Berlin Heidelberg New York (2004) 761 - 769
5. Prentke Romish Company: http://store.prentrom.com/cgi-bin/store/TR2000.html
6. Evans, D.G., Drew, R., Blenkhorn, P.: Controlling mouse pointer position using an infrared head-operated joystick. IEEE Transactions on Rehabilitation Engineering, Vol. 8. IEEE Engineering in Medicine and Biology Society (2000) 107-117
7. Pregenzer, M., Pfurtscheller, G.: Frequency component selection for an EEG-based brain to computer interface. IEEE Transactions on Rehabilitation Engineering, Vol. 7. IEEE Engineering in Medicine and Biology Society (1999) 413-419
8. Chen, Y.L.: Application of Tilt Sensors in Human-Computer Mouse Interface for People With Disabilities, IEEE Transactions on Rehabilitation Engineering, Vol. 9. IEEE Engineering in Medicine and Biology Society (2001) 289-294
9. Kim, Y.W.: Development of headset-type computer mouse using gyro sensors for the handicapped. IEE Electronics Letters, Vol. 38. (2002) 1313- 1314
10. Rao, R.S., Seliktar, R., Rahman, T.: Evaluation of an Isometric and a Position Joystick in a Target Acquisition Task for Individuals with Cerebral Palsy. IEEE Transactions on Rehabilitation Engineering, Vol. 8. IEEE Engineering in Medicine and Biology Society (2000) 118-125
11. Surdilovic, T., Zhang, Y.Q.: Convenient Intelligent Cursor Control Web Systems for Internet Users with Severe Motor-Impairments. International Journal of Medical Informatics. Available Online 2005
12. Trewin, S., Pain, H.: Keyboard and mouse errors due to motor disabilities. International Journal of Human-Computer Studies, Vol. 50. (1999) 109–144
13. Bluetooth SIG: Bluetooth Specification Version 2.0+EDR. Bluetooth SIG Standard (2004)
14. Percept Technology Report: http://download.microsoft.com/download/1/3/9/139a8c30-34cc-4453-a449-7a1c586a3ae5/MicrosoftMouseBatteryExecSummary.pdf

Smart Wheelchair Based on Ultrasonic Positioning System

HyungJun Kim and DaeHyun Ryu

Division of Information Technology
Hansei University, Korea
harry@hansei.ac.kr

Abstract. We present a robot wheelchair based on a UPS (Ultrasound Positioning System) for the elderly or the disabled. The proposed robot wheelchair system, called Smart Wheelchair, is a standard powered wheelchair equipped with a laptop computer which supports a variety of GUIs (graphical user inter-faces) for the disabled and a microprocessor system equipped with UPS and external connectivity modules. The Smart Wheelchair incorporates state-of-the-art technology providing information about the location of the wheelchair based on its movement, which is beyond comparison to ordinary wheelchairs that have just one function: to move patients. The Smart Wheelchair can provide users with driving assistance and take over low-level navigation to allow the user to travel efficiently, safely and easily.

1 Introduction

Our research project is aimed towards developing a usable, low-cost assistive smart wheelchair system for the disabled. We have developed an indoor navigation system based on UPS (Ultrasound Positioning System) and a graphical user interface, so that disabled or senior people can navigate autonomously using the Smart Wheelchair. Many people in wheelchairs are unable to control a powered wheelchair with the standard joystick interface. The Smart Wheelchair takes commands from the user based on a simple graphical interface of computer and executes them while taking into consideration local positioning information. This will allow the chair to follow preset paths and avoid obstacles. It is very important to find the local position in mobile robot navigation.

The smart wheelchair is equipped with a ATmega128 processor for the positioning and controller module, which can recognize the position of the wheelchair using information from U-SAT system, and control the motor. By using a laptop computer or PDA, the user can send commands to the control system. Depending on different user needs, various HCI (human computer interfaces) options may be needed to provide control. The Smart Wheelchair can provide users with driving assistance and take over low-level navigation to allow the user to travel efficiently, safely and easily.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 1014–1020, 2006.

2 Related Works: Robot Positioning

The robot positioning system has been studied and researched for many years. Robot positioning can be categorized into three parts: absolute, relative positioning, and a combination of both methods. The dead-reckoning method is one of the most popular among relative positioning methods. The dead-reckoning method uses encoded information which it gains from the wheels to determine the position of the robot. But because of wheel slippage, mechanical tolerance and surface roughness, the accumulation of errors for this method is unbounded. The distance between the actual position and the computed position increases as it moves further. So the actual position is hardly maintained as it moves over longer distances [1]. Other methods use sensors such as a rate-gyro and a magnetic compass. A rate-gyro accumulates errors continuously with the passage of time and a magnetic compass does not function well in a place where magnetic fields vary from position to position. For these reasons, it is hard to find accurately the location of the robot moving over long distances by only using relative positioning.

Global Positioning System (GPS) is mainly used for absolute positioning. Today's GPS receivers are extremely accurate, thanks to their parallel multi-channel design. Certain atmospheric factors and other sources of error can affect the accuracy of GPS receivers. Garmin GPS receivers are accurate to within 15 meters on average. Newer Garmin GPS receivers with WAAS (Wide Area Augmentation System) capability can improve accuracy to less than three meters on average. Users can also get better accuracy with Differential GPS (DGPS), which corrects GPS signals to within an average of three to five meters. But three satellites are necessary to get a three dimensional coordinate of the GPS receiver attached on a robot. In addition one more satellite is necessary because of clock bias. Therefore it does not operate well in places where no GPS satellite signal reaches or there are less than four visible satellites [2]. The absolute position of a robot can be measured by using GPS, however, it is not accurately measured in places where the GPS satellite signal is blocked.

The methods which measure the absolute position of a robot without using GPS are CCD cameras, infrared light and ultrasonic sensors. First, the method of using a CCD camera requires complicated signal processing to analyze images. In addition it is expensive and depends highly on camera calibration and image sensitivity [3,4]. Second, finding the absolute position of a robot by using infrared light is easy and inexpensive as the incident angles of the infrared light can be detected by using an incident angle sensor. However, this method leads to low system performance, hardly applicable to an outdoor environment [5]. Third, there are two methods of detecting absolute position using ultrasonic sensors. One method uses the reflection of ultrasonic waves and another uses ultrasonic waves directly. The method using the reflection of ultrasonic waves is useful for the detection of obstacles but is inappropriate for the determination of absolute position. Therefore the measurement of distance using direct ultrasonic waves has generally been used. The method using direct ultrasonic waves shows better performances and is more economical compared to the other methods referred to above. In this method, an ultrasonic transmitter is located on a robot and several ultrasonic receivers are attached to the ceiling to find the position of the robot [6]. Ultrasonic transmitters are also located at fixed positions whose coordinates are known and are read by an ultrasonic receiver attached to the robot [7,8]. In the measurement of the distance using direct ultrasonic waves, a high-precision method is proposed [9]. The absolute positioning system using ultrasonic sensors based on this method is represented as UPS. UPS is an absolute positioning system using direct ultrasonic waves and has better performance and is more economical than other absolute positioning systems.

3 System Design and Implementation

3.1 Ultrasonic Positioning

Fig. 1 shows the concept of ultrasonic positioning system in the SmartWheelchair positioning, which is constructed by multi-ultrasonic-transmitters placed in the environment and one receiver mounted on the wheelchair. This system uses direct wave reception rather than reflected wave reception. Not only 2-dimensional but also 3-dimensional position information can be determined using the U-SAT (ultrasonic satellite) system. The U-SAT is an ultrasonic positioning system that can obtain an absolute position using known positions of ultrasonic transmitters and a receiver attached to a powered wheelchair. The advantage of U-SAT is that the position of the wheelchair is determined easily and economically.

Ultrasonic transmitters with a RF (radio frequency) synchronous broadcaster and a receiver are used to measure distance as shown in Fig. 2. The ultrasonic transmitters and receiver of this precision distance measuring module utilize a

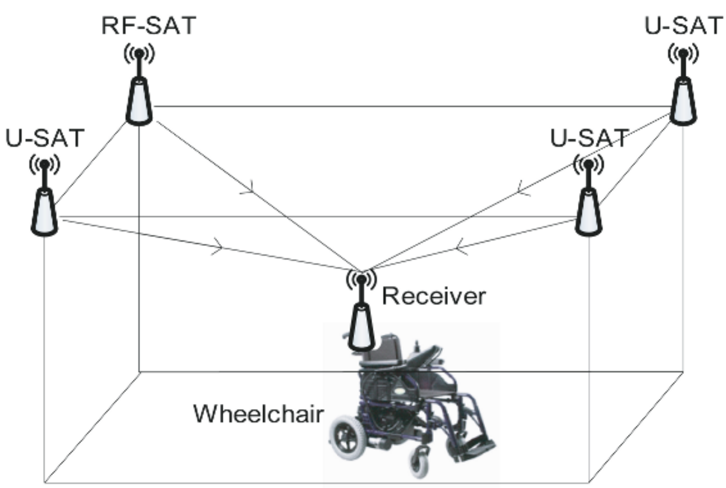


Fig. 1. U-SAT system for the Smart Wheelchair positioning

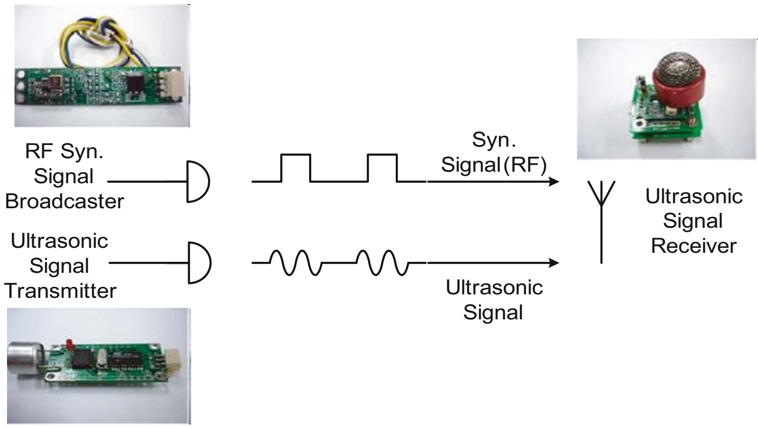


Fig. 2. RF-broadcaster, U-SAT transmitter, and U-SAT receiver

frequency of 40 KHz. When the ultrasonic transmitters receive a synchronous signal from the RF-broadcaster, the transmitters start to generate ultrasonic signals at 40 KHz. The receiver checks the time difference between the RF-broadcaster's signal and the ultrasonic transmitters' signals, and then determines its position using TOF (Time of Flight). In this experiment the distance limit is about 20m and measurement errors are less than 1cm.

3.2 Smart Wheelchair System Design

The general configuration of the Smart Wheelchair system is shown in Fig. 3. The Smart Wheelchair is equipped with a ATmega128 processor for the positioning and controller module. The positioning and controller module can recognize the position of the wheelchair using information from U-SAT system, and control the motor. It also can get information from close-in sensors and acceleration sensors. By using a laptop computer or PDA, the user can send commands to the control system. Depending on different user needs, various HCI (human computer interfaces) options may be needed to provide control based on voice recognition, gesture, or motion.

The system can communicate with a server using a short distance communication system such as WiFi (Wireless Fidelity), Bluetooth, or Zigbee. The server is designed not only to follow the location of the Smart Wheelchair but also to monitor the condition of its users in real time. This system will make use of various medical sensors which will interface with the communication system. The user's condition may be monitored from a distance by a medical doctor or family members through the Internet. The proposed system can be adapted to users depending on the user's degree of disability. We have configured the positioning and navigation system for an indoor environment for our experiments. A laptop PC is used for the robot's graphical user interface. The focus was on creating an interface that could be easily customized for various users and their access methods.

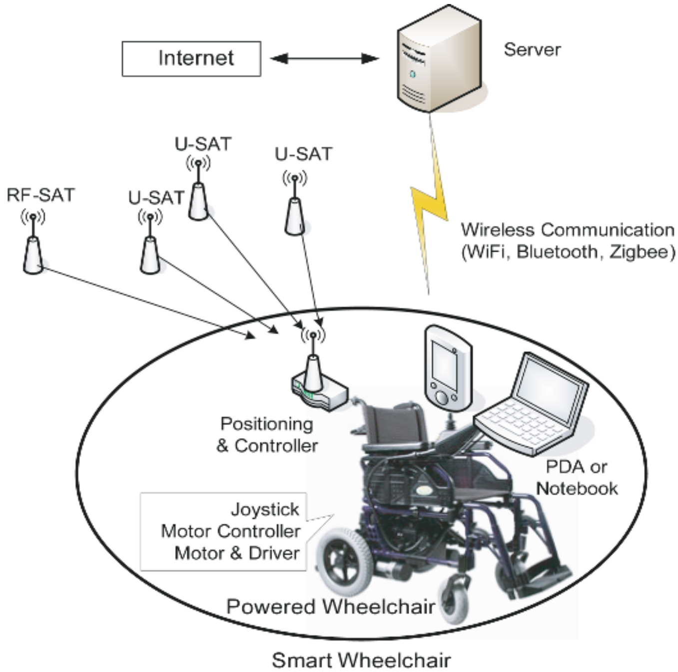


Fig. 3. Configuration of the Smart Wheelchair system

3.3 User Interface

Fig. 4 shows an example of the User Interface (UI). There is a small window to show the trajectory of the path on the top left corner of the UI. An example of the implementation of path tracking is shown in Fig. 5.

4 Conclusions and Future Works

This research project is aimed towards the development of a usable, low-cost assistive smart wheelchair system for the disabled. In the initial work towards this goal, an indoor navigation system based on UPS (Ultrasound Positioning System) and a graphical user interface have been developed. The Smart Wheelchair must work with the user to accomplish the user's goals, accepting input as the task progresses, while preventing damage to the user, other obstacles and the Smart Wheelchair itself.

Work is continuing towards the goal of a complete smart wheelchair system. A robotic wheelchair must be able to navigate in both indoor and outdoor environments. While indoor navigation can implemented using infrared and sonar sensors, outdoor navigation can not rely on these sensors alone. The Smart Wheelchair will continue to take high-level directional commands from the user and execute them while keeping the user safe. The system will automatically select indoor or outdoor mode using an indoor/outdoor sensor currently in development.

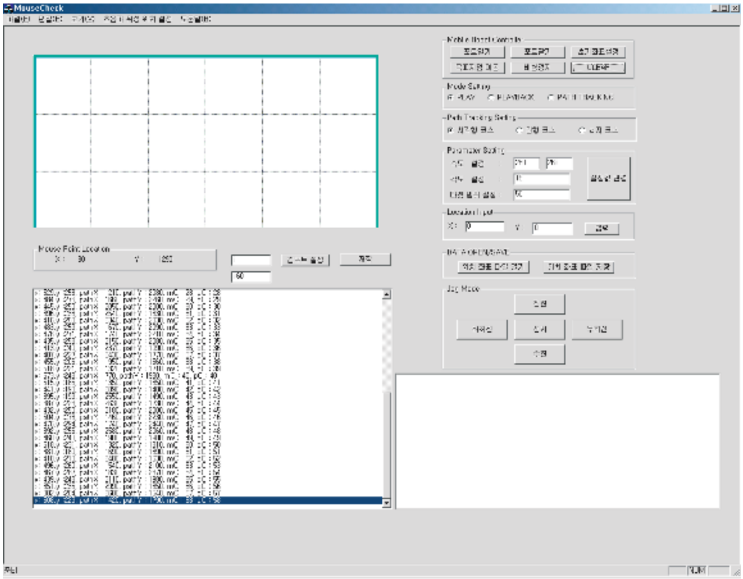


Fig. 4. Example of UI for local positioning and control

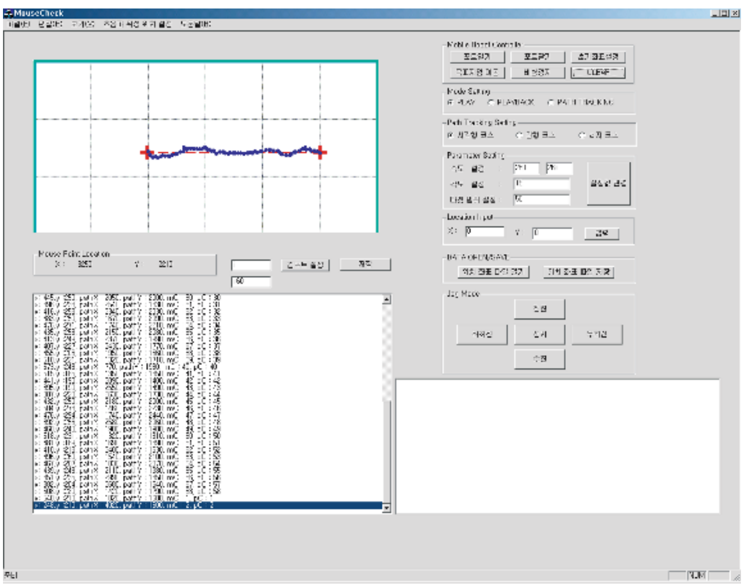


Fig. 5. Example of UI for wheelchair tracing

Acknowledgements

This work was supported by Hansei University.

References

1. Kim, J. H., Seong, P. H.: Experiments on Orientation Recovery and Steering of Autonomous Mobile Robot Using Encoded Magnetic Compass Disc. IEEE Trans. Instrumentation and Measurement 45(2) (1996) 271–274.
2. Evans, A. G., Swift, E. R., Cunningham, J. P., Hill, R. W., Blewitt, G., Yunck, T. P., Lichten, S. M., Ron Hatch, R., Malys, S., Bossler, J.: The Global Positioning System Geodesy Odyssey Navigation. Journal of The Institute of Navigation 49(1) 2002 7–33
3. Wijesoma, W. S., Kodagoda, K. R. S., Balasuriya, A. P.: A laser and a Camera for Mobile Robot Navigation. in 7th Int. Conf. on, Control, Automation, Robotics and Vision (2002) 740–745
4. Song K., Tang, W.: Environment Perception for a Mobile Robot Using Double Ultrasonic Sensors and a CCD Camera. IEEE Trans. Industrial Electronics 43(6) (1996) 372–379
5. Arai, Y., Sekiai, M: Absolute Position Measurement System for Mobile Robot based on Incident Angle Detection of Infrared Light. in Int. Conf. Intelligent Robots and Systems (2003) 986 – 991
6. Ghidary, S. S., Tani, T., Takamori, T., Hattori, M.: A New Home Robot Positioning System (HRPS) Using IR Switched Multi Ultrasonic Sensors. in IEEE Int. Conf. Systems, Man, and Cybernetics (1999) 737–741
7. Tsai, C. C.: A Localization System of a Mobile Robot by Fusing Dead-reckoning and Ultrasonic Measurements. IEEE Trans. Instrumentation and Measurement 47(10) (1998) 1399–1404
8. Yi, S. Y., Jin, J. H.: Self-localization of a Mobile Robot Using Global Ultrasonic Sensor System. Journal of Control, Automation, and Systems Engineering 9(2) (2003) 145–151
9. Lee, D. H., An, H. T., Baek, K. R., Lee, M. H.: A study on Resolution Enhancement in the Detection of Ultrasonic Signal. in The KSAE Annual Fall Conference (2003) 95–102

Scenarios of Use for a Modular Robotic Mobility Enhancement System for Profoundly Disabled Children in an Educational and Institutional Care Environment

Peter Mayer[1], Paul Panek[1], Georg Edelmayer[1], Marnix Nuttin[2], and Wolfgang Zagler[1]

[1] fortec - Research Group on Rehabilitation Technology,
Vienna Univ. of Technology, Favoritenstrasse 11/029, A-1040 Vienna, Austria
{mayer, panek, ged, zw}@fortec.tuwien.ac.at
http://www.fortec.tuwien.ac.at/movement
[2] Katholieke Universiteit Leuven, PMA
http://people.mech.kuleuven.be/~mnuttin

Abstract. In the framework of the EU funded MOVEMENT Project a novel modular robotic system is being developed which aims at supporting the mobility of elderly citizens and persons with disabilities. This paper outlines some of the use cases which were developed for the envisaged system with focus on assisting severely disabled children and their carers in an institutional environment. Six scenarios were developed and commented by nine professional carers. The paper presents and discusses the qualitative and quantitative data gained from user panel discussion. It was found that the concept and the use cases of MOVEMENT system were very well rated regarding utility which confirms the work of the consortium up to now. The future development activities towards the roll out and evaluation of the first MOVEMENT prototype platform are described.

1 Introduction and Aim

MOVEMENT stands for "Modular Versatile Mobility Enhancement Technology" and aims at the development of a novel system for supporting mobility of disabled and elderly persons. The core of the envisaged MOVEMENT system is formed by an intelligent mobile (robotic) platform which can attach to a user definable selection of application modules (e.g. chair, manipulator, information terminal). These modules are more or less inconspicuous mainstream articles but will become powerful assistive devices when the mobile platform attaches to them [12]. Due to the novel approach of MOVEMENT the initial development and refinement of use cases were crucial in order to present the concept to the future users and to gain feedback from them. While the consortium developed scenarios for a wide variety of environments (private home, institutional care, shops, sheltered living) and different user groups (persons with mild/temporary, moderate and severe disabilities) [5] the paper on hand focuses on a subset thereof, namely on those scenarios which describe the future application of MOVEMENT for severely disabled children and adolescents at a support centre.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 1021 – 1028, 2006.

2 State of the Art

Several RTD projects and research groups are working on the development of smart wheelchairs [1,2,6,7,8,9,11,14,15]. Thus, the last decade saw the evolution of more and more complex wheelchairs demonstrating capabilities for navigation, manipulation and transport. Unfortunately, commercialisation usually was difficult or even not possible, since the prototypes are bulky, difficult to operate, need to be configured for each individual person and are still very expensive.

3 Vision and Concept of MOVEMENT

MOVEMENT aims at developing a new solution for supporting personal mobility which meets the users' expectations for an inconspicuous, non-stigmatising, tailorable, ready to use and affordable mobility aid. As a consequence, the objectives of the project are:

- Addressing all three aspects of mobility (moving people, objects and information) by a fully modular set of assistive devices that can be freely assembled depending on the user's needs.
- Providing a concrete solution which can be placed on the assistive technology market soon after completion of the project.
- Pursuing an active dissemination and demonstration strategy by which users, caregivers and the health system are informed about the product under development, leading to awareness creation on a European level.

MOVEMENT addresses the needs of persons that do not need or want to use a wheelchair all the time but just intent to use it occasionally for managing their way from room to room. Persons normally not able to control a wheelchair (e.g. in case of spasm or athetosis) will also profit from the system developed within the MOVEMENT project. It, therefore, will be possible to drive the MOVEMENT system in a manual, an assisted or an autonomous mode [4]. To support the project goals, the User Interface for MOVEMENT will consist of a modular combination of traditional input/output with additional input channels and feedback, adaptive behaviour, prediction of user intention in combination with shared control, autonomous navigation and collision avoidance. For all operating modes the User Interface will adapt to the user characteristics and specific habits so that the initial setup will be improved over time. Main innovations the project aims at are:

- Modularity to enhance the flexibility of use and to reduce costs
- High level commands and shared control principle
- Enhanced flexibility due to modularity
- Mobile, multi-functional communication terminal approaching the user on demand.

4 Methodological Approach for User Needs Investigation

Detailed scenarios and clear pictures regarding MOVEMENT modules have been developed ensuring an easy to understand presentation to the users and enabling them to provide clear feedback and comments. Main purpose was to collect primary and secondary users' comments, needs and wishes regarding the foreseeable application of the MOVEMENT modules in daily usage. This was done by telling stories, by creating separate scenarios for different settings even if from technical point of view the scenarios are similar. These scenarios were developed and pre-tested in direct cooperation with some key users and experts. The focus is on system's functionality and the benefits for the users. Totally the consortium developed a comprehensive set of 35 scenarios which were presented to 45 participants of the user panels organised in Italy, Austria and The Netherlands [3]. Note that this paper focuses on 6 out of the 35 scenarios. For the user panels the following procedure was agreed on:

1. explanation of the purpose of the meeting based on slides
2. explanation of the basic MOVEMENT concepts based on slides
3. oral presentation of pre-selected scenarios supported by slides (handouts)
4. discussion on each of the scenarios on the following questions
5. could you imagine *yourself* using this solution? (Y/N) [if not, why not?]
6. could you imagine *someone else* using this solution? (Y/N)
7. can you foresee use in other use conditions (free text)
8. can you foresee barriers/problems during usage? (free text)
9. which priority would you give to making this scenario come true *for you in your specific situation*? (very high / high / medium / low / very low)
10. which priority would you give to making this scenario come true *for disabled and elderly in general*? (very high / high / medium / low / very low)
11. discussion on each scenario; comments, evaluation, suggestions?

For each scenario quantitative and qualitative data were gained: The quantitative results were regarding utility (yes =100%, no=0%) and priority (high=100%, high=75%, medium=50%, low=35%, very low=0%). The qualitative results (quoting of the oral and/or written comments) were assigned as far as possible to the user code of the originator in order to allow back tracing.

In the beginning of the meetings the participants received and signed an informed consent form [10]. This form contains information about the MOVEMENT project, the concrete objective of the user panel activities and the activities to be done during interview. It also names the responsible researcher with contact information and makes clear that the participation in the project is completely on a volunteer base and can be withdrawn by the participant at any time without any justification. It also explains what the collected data are used for and that the data will be coded and only used for scientific reasons inside the consortium.

The consortium communicated clearly that not all scenarios will be implemented in the framework of the project (even if there might be high need from the users' side) due to limited resources.

5 Scenarios of Use

Six out of the 35 scenarios [3] are showing usage of future MOVEMENT system with regard to severely disabled children and youngsters and their carers. Description of environment: A support centre for motor and multiple impaired children and adolescents (aged 5 - 18) near Innsbruck, Austria. It runs a kindergarten, school and residential school. There are 104 children, about 20 of them could benefit from MOVEMENT.

5.1 Transferring Children in Standard Chairs Between Rooms in School Building

10 o' clock. The bell indicates that the lesson is closed. The next lesson will take place in another room (physiotherapy). Some of the children in the class room are using their UI to call the platform. When the platform has arrived behind the child's standard chair (of type "Tripp Trapp") it makes a beep to indicate its readiness for docking. The child confirms via the UI. The platform now slowly moves under the chair, stops there and starts to move the chair up for about 5 cm. The child is informed via speech output that the chair has been docked and the platform now is ready to go. The child uses the joystick to steer the platform (manual or assisted driving) and moves into the therapy room. There the child steers the platform into a position which is used for the later manual transfer out of the chair which will be assisted by the physiotherapist. The child uses the UI to activate the undocking process. Thus, the platform moves down the chair about 5 cm till the chair has stable contact with floor. Then the platform moves backwards a bit. It informs the child that undocking was completed successfully and leaves the room to travel to the next child waiting for a lift.

5.2 Assisted Driving

Charles is sitting in his orthopaedic chair which is docked on the platform. He is a young boy and not yet familiar with driving a wheelchair. Thus, when moving around a corner it often happens that he underestimates the space needed between the chair and the wall. In such situations the platform stops automatically in order to avoid a collision between the platform and the vertical edge of the wall which Charles is coming around. After the platform has stopped Charles can steer another course. The assisted driving feature is also providing support to Charles when he is steering the platform through a narrow door. When ever a collision is foreseeable the platform automatically reduces speed and stops automatically when necessary. The platform also recognises stairs and stops to avoid a fall downstairs.

5.3 Smart Wheelchair for Transferring Mail

Tim is quite proud of his new daily duty. Some days ago the head of school asked him if he would like to do him a favour: to bring the mail of today into the head office. Everyday in the beginning of the first break Tim moves with his platform to the entrance hall asking the porter for the mail. The porter hangs the bag with the mail on the platform and Tim moves through the corridor to the head office of the school. Usually the door is open and he calls the director from outside who then will take up the bag with the mail. Today the door is closed. Obviously the director is not here. Tim is not happy about this as he enjoys a lot the small conversations which usually are taking place between him and the director. In order to complete his job he is moving the platform next to the door in parallel to the wall and is activating the object transfer mechanism. This makes the platform hang the bag with the mail onto a hook on the wall next to the closed door. Then Tim moves away to use the remaining part of the break for meeting his friends.

5.4 Special Chair for Severely Disabled Children

Most time Sue is lying in a horizontal position. For medical reasons several times a day she is brought into a vertical position. Since she has the M-system her mobility has increased significantly compared to the previous situation. The mobile platform can dock to a special chair which can change its own shape from a bed like shape into a vertical position. With this system, which is steered by the care persons, it is much easier for Sandra to participate in activities of the other children. The care persons appreciate the reduction of their own physical load.

5.5 Mobile Board for Communication, Information and Environment Control

Over the week Tom is living together with 7 other youngsters in one of the groups of a residential school for disabled persons. In the morning he attends school in the same institution, afternoon and night he spends at the residential school. During weekend he usually is with his parents in a village 50 km away. But this weekend he will stay at the institution. For Saturday evening he plans to watch a movie. First, he wants to call up one of his friends. He uses the M- user interface mounted on his chair to call the table with the mobile information board. The platform automatically moves to the room where the information board currently is located, docks it automatically and moves with it in front of Tom. Tom now can steer the information board via his MOVEMENT user interface. He activates the internet access and checks the starting time of a movie which he wants to watch on TV. After he has found this information he activates the phone mode and selects the numbers of his friend he wants to call to. After the farewell he hangs up the phone and activates the environmental control system on the terminal to activate the TV and to select the channel where the movie will be delivered. He is checking the volume level of the TV and then he decreases the light level of the room light. After this he tells the platform to move the terminal away. While Tom is enjoying the first sequence of the movie the M-platform quietly is moving the terminal out of the room.

5.6 Mobile Board for Environment Control

It's 10 to 8 in the morning. Ellen who is main teacher in one of the classes has already completed her preparation, all the children have arrived. As usual, the first lesson will start with the "Morgenkreis" (morning circle) which is a ritual nearly all of the pupil do like a lot. The children come together, sitting down on the floor or sitting in their chairs and wheel chairs in the shape of a circle. Ellen invites them to share their impressions from yesterday and to have a look on what this day will bring. The ritual is completed with a song, today it is a song from a CD-Player. Ann who is using a M-system compatible tripp trapp calls the platform which is carefully moving directly behind the tripp trapp. It signals its readiness for docking. Ann confirms via user interface and the tripp trapp is docked. Ann steers the system towards her position in the circle where she undocks. She then calls the platform to bring the environmental control terminal next to here. The platform automatically moves in the corner of the class room, docks the terminal, moves the terminal next to Ann. Ann uses the UI to control the terminal and selects the environmental control feature which enables here to start the CD-Rom with the song on it.

6 Findings

Nine professional assistants (teachers, therapists) of severely disabled children and youngsters participated in the evaluation of 6 scenarios. The scenarios were accepted very well, the users saw them as very good examples of what the platform should actually be able to do.

Not surprisingly, one of the major findings is the explicit mentioning of safety related issues. The users must have the final control over the platform in any case. There have to be precautions that the platform does not e.g. go down stairs (automatically and by users' intent). Speech input is seen as an interesting feature, however not all users (especially of the 'severe' group) are capable of producing speech in a quality that speech recognition software would need. Well designed alternative input methods are therefore necessary. It turned out, that sharing of a platform between multiple users (as presented in some scenarios) is seen quite interesting by the users (children's home, sheltered home) , especially as this might limit costs of the total system. For safety while driving it is important that the platform reacts on the movement of the child's arms. If the arms are stretched left and right of the platform this most likely will exceed the geometric width of the platform. This must be checked by the platform before moving through a narrow door. For recognising obstacles one should be aware that there might be several things on the floor which should not cause the system to stop (e.g. paper, cloth, threads) but also obstacles with nearly the same height (a hand of a child) which must be recognised as an obstacle causing the system to stop.

A transport box with a magnetic mechanism to dock / undock to platform (or to conventional power wheelchair) would be very useful as general way for transport of boxes which can be docked and undocked by the user autonomously. The boxes can be filled and purged independently. MOVEMENT also was seen as a learning system which assists users to learn driving a mobility system [13].

Table 1. Ranking of utility and priority of 6 out of 35 scenarios as seen by 9 professional carers of a support centre for motor and multiple impaired children and youngsters in Austria

Scenario	involved MOVEMENT modules	utility	(answers)	priority	(answers)
SC1 - Transferring Children in standard chairs between rooms in school building	platform, table	62,50	8	60,00	5
SC2 - Assisted Driving	platform, orthopedic chair	100,00	9	82,22	9
SC3 - Smart Wheelchair for Transferring Mail	platform, chair, container	100,00	9	75,56	9
SC4 - Special Chair for severely disabled children	platform, special chair	71,43	7	60,00	5
SC5 - Mobile Board for Communication, Information and Environment Control	platform, chair, ICT terminal	100,00	6	93,33	6
SC6 - Mobile Board for Environment Control	platform, chair, ICT terminal	100,00	6	76,67	6

Not all users of the panel have evaluated all scenarios, this fact has to be taken into account when interpreting quantitative data however, and information gained from the qualitative input covers all the scenarios. Some scenarios present the same 'feature' of the MOVEMENT system but in different settings or surrounding for the different user groups. Although a manipulator is not part of the MOVEMENT system and was not presented in the scenarios, some users suggested (or asked for) a possibility to pick up items from the floor or to have support when one wants to take something from a shelf of a general form of handling support.

The presented scenarios of use were well accepted by severely disabled users and by care persons. Valuable new feedback from users was collected and systematically documented and is used as input for the ongoing development process. This detailed knowledge about user needs, helps the consortium in carrying out the research and development activities in a user centered way while focusing on those modules which actually will be implemented in the framework of the project. Generally (see also [3]) it was found that the concept and the use cases of MOVEMENT system were very well rated regarding utility which confirms the work of the consortium up to now.

The aspects of running the MOVEMENT system in institutional environments allowed gaining some valuable information regarding multi user usage and regarding economic aspects (one platform could serve several users; the platform could be part of the infrastructure and therefore could be paid by the institution).

Acknowledgements

MOVEMENT (contract number 511670) is co-funded by the European Commission. Project partners: Vienna Univ. of Technology, ARC Seibersdorf Research GmbH, iRv - Institute for Rehabilitation Research, BlueBotics SA, Otto Bock Health Care, Technische Universität München, Katholieke Universiteit Leuven, Scuola Superiore Sant' Anna. Regarding the user needs investigation reported here we are thankful for valuable input delivered by Support Centre Elisabethinum Axams, Tyrol, Austria.

References

1. Borgolte U., Hoyer H., Bühler C., Heck H., Hoelper R.: Architectural Concepts of a Semi-autonomous Wheelchair, Intelligent and Robotic Systems, vol. 22, 1998, pp. 233-353
2. Bühler C., Heck H., Humann W.: User-Driven Human-Machine Interface Configuration for a Wheelchair with Complex Functionality, in: Anogianakis G., Bühler C., Soede M. (eds.): "Advancement of Assistive Technology", Proc of AAATE , IOS Press, 1997, pp. 375-380
3. D1.1a - Report on User Abilities, Preferences and Needs, MOVEMENT Project Deliverable, to be published mid of 2006 on website of consortium [12]
4. G. Edelmayer, H. Van Brussel, E. Demeester, A. Hüntemann, G. Kronreif, P. Mayer, G. Vanacker, D. Vanhooydonck, W.L. Zagler, M. Nuttin: "Enhanced User Interface for Power Wheelchairs Developed in the MOVEMENT Project"; EMBEC 05 3rd Europ Medical & Biolog Eng Conf, Prague; in: "IFMBE Proceedings", 11 (2005), ISSN 1727-1983; 4 S.

5. G.J. Gelderblom, G. Kronreif, A Pisetti, P. Mayer: "User need assessment in MOVEMENT"; 3rd CWUAAT, Cambridge, 2006
6. Katevas N.I., Sgouros N.M., Tzafestas S.G., et al: The Autonomous Mobile Robot SENARIO: A Sensor-Aided Intelligent Navigation System for Powered Wheelchairs, IEEE Robotics Automation Magazine, December 1997, pp. 60-70
7. Lankenau Axel, Avoiding Mode Confusion in Service Robots – The Bremen Autonomous Wheelchair as an Example, in: Mokhtari M. (ed.): Integration of Assistive Technology in the Information Age, Conf Rehabilitation Robotics, IOS Press, 2001, pp.162-167
8. Lankenau, A. and Rofer, T. (2001). A safe and versatile mobility assistant. IEEE Robotics and Automation Magazine, 8(1):29–37.
9. Levine Simon P., Bell David A., et al., The NavChair Assistive Wheelchair Navigation System, IEEE Trans. on Rehab. Engineering, vol. 7, no. 4, pp. 443-451, Dec. 1999
10. M. Rauhala, I. Wagner: Ethical Review - A Continuous Process in an Assistive Technology Project, Proc. AAATE conf., IOS Press, 16 (2005), ISBN 1-58603-543-6; 31 - 35.
11. Matsumoto Yoshio, Ino Tomoyuki, Ogasawara Tsukasa, Development of Intelligent Wheelchair System with Face and Gaze Based Interface, Proceedings of the 10th IEEE International Workshop on Robot and Human Communication (ROMAN 2001), pp. 262-267
12. MOVEMENT project homepage `http://www.fortec.tuwien.ac.at/movement`
13. Phil Odor and Marion Watson: Learning through Smart Wheelchairs, Call Centre, Univ. of Edinburgh, 1994
14. Pollack M.E., Engberg S., Matthews J.T., Thrun S., Brown L. et al., Pearl: A Mobile Robotic Assistant for the Elderly, AAAI Workshop on Automation as Eldercare, 2002
15. Sensor-Assisted Wheelchair Project
    `http://www.mech.kuleuven.be/pma/project/stww`
    `http://www.mech.kuleuven.be/mlr`

Blind and Visually Impaired People: Human Computer Interface
Introduction to the Special Thematic Session

Grigori Evreinov

Dept. of Computer Sciences, University of Tampere
Grigori.Evreinov@cs.uta.fi

For over ten years human-computer interface, blind interaction and integration of visually impaired users with sighted users are the key issues of equal access to information and service. The vast research on alternative visualization, augmented communication, user-centered design and usability has been done, and much more projects and solutions are under development. However, several generations of graphical interfaces (Xerox, Apple, Microsoft) have brought less or no benefits for the blind users. Some elderly people still recall the times of DOS and command line, when both the system and application software levels were almost equally accessible. Nowadays, multi-processor operating systems are extremely complex and perform hundreds of routine tasks which are not necessary to be supervised or adapted for the user control at all.

Of course, we cannot diminish advances in speech and dialog processing. Nevertheless in practice, at the level of “terminal window” we cannot observe a great progress that has been done during the last years to provide easy access for blind users. A huge number of software, hardware prototypes and theoretical models have been developed for transforming user interface or components to be unified and adaptable for all [8], [9], [10]. And yet a good interface seems to be more practical and robust when different concepts and metaphors might especially be modified for the target group of users. While, the same technique may not be suitable and efficient for all, e.g., Braille interface [6] and multimodal integration myths [7].

Graphical input and visualization have irrefutable benefits as an intermediate between an analog perception of the human being and a digital “assistant processor”. Unfortunately, ten years in the field of auditory display research did not lead to a significant progress in visualization of sound images [4]. Other techniques making use of electro- or vibro-tactile imaging have a longer history. Up to now only a simulation of simple geometrical shapes may evoke sensations similar to their visual prototypes [1], [3], [5]. Diverse pin-matrixes with a low usability have been produced as an imitation of the visual display principles that are often not acceptable for cognitive processing a dynamical tactile image [6].

“GUIs are unsuitable for blind use” – we can find such a statement in almost each paper where authors carried out usability study of some improvements or have proposed an alternative approach and techniques [2], [12]. Still, some authors try to improve the methods and add more features which can bring frustration rather than benefits for non-visual communication and an access to semantics of the message

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 1029 – 1030, 2006.

encoded [11]. An access to information content in its various form is the goal of the end user. Navigation, pointing and selection are not the goal, but just one of the ways which may be appropriate or unsuitable being used with the particular methods, tools, environment and the task.

It is a great challenge to develop human-computer interfaces which can be properly matched to the user needs and abilities. To support blind interaction with entities of the information space new techniques ought to be based on metaphors and paradigms which are inherent and intuitive for the blind computer user. Blind people need a good interface design that will obviously differ from the existing solutions, which have been designed and intended, for instance, for visually impaired people.

We make a lot of errors and create the prototypes which probably will never be used. Sometimes, we invent algorithms and interaction models which contradict to the nature of the human perception. However, we know the goal and we are looking for a solution. Our knowledge and experience in developing the means of information delivering and manipulation will grow up. And we cannot say that nowadays visually impaired people just have an access to Braille sheets and use a white cane.

In this session the authors will discuss the issues and prototypes which sometimes would become a vital component, tool or technology making a natural integration of the blind users in high-tech society.

References

1. Bach-y-Rita P, Collins CC, Saunders F, White B, Scadden L. Vision substitution by tactile image projection. Nature 221, (1969) 963-964
2. Boyd, L.H., Boyd, W.L. and Vanderheiden, G.C. The graphical user interface: Crisis, danger and opportunity. In J. of Visual Impairment and Blindness, Dec. (1990) 496-502
3. Hiroaki Takahashi, et al. Electro-Tactile Display with Localized High-Speed Switching. In Proc. of ICAT 2002, Tokyo, Japan (2002) 10-16
4. International Community for Auditory Display. Website 2006 `http://www.icad.org/`
5. Kajimoto, H et al SmartTouch - Augmentation of Skin Sensation with Electrocutaneous Display, In Proc. of HAPTICS 2003, (2003) 40-46
6. Lévesque, V., Pasquero, J., Hayward, V., Legault, M. Display of virtual Braille dots by lateral skin deformation: feasibility study. TAP 2(2), (2005) 132-149
7. Oviatt, Sh. Ten myths of multimodal interaction. Communications of the ACM 42, (1999) 74-81
8. Richter, K. and Enge, M. Multi-modal Framework to support users with special needs in interaction with public information systems. In Proc. of Mobile HCI, Udine, Italy, Springer, LNCS, 2795, (2003) 286-301
9. Savidis, A., Stephanidis, C. Developing Dual Interfaces for Integrating Blind and Sighted Users: The HOMER UIMS. CHI 1995, (1995) 106-113
10. Savidis, A., Stephanidis, C. Unified user interface design: designing universally accessible interactions. Interacting with Computers 16(2): 243-270 (2004)
11. Tang,H. and Beebe,D.J. An Oral Tactile Interface for Blind Navigation. IEEE Trans. on Neural Systems and Rehab. Engineering, 14(1), (2006) 116-123
12. Wiecha, Ch., Szekely, P. Transforming the UI for anyone, anywhere. Enabling an increased variety of users, devices, and tasks through interface transformations. In Proc. of CHI 2001. Workshop. (2001) 493-484

A Framework for Blind User Interfacing

Fernando Alonso, José L. Fuertes, Ángel L. González, and Loïc Martínez

Facultad de Informática, Universidad Politécnica de Madrid, Campus de Montegancedo, 28660 Boadilla del Monte, Madrid, Spain
{falonso, jfuertes, agonzalez, loic}@fi.upm.es

Abstract. There are specific usability requirements that have to be met when developing dual interfaces, that is, graphical user interfaces that are adapted for blind users. These include task adequacy, dimensional trade-off, behavior equivalence, semantic loss avoidance and device independence. Consequently, the development of human-computer interfaces that are based on the task, domain, dialog, presentation, platform and user models has to be modified to take into account these requirements. This paper presents a framework that includes these requirements, allowing for the development of dual interfaces. The framework includes a set of guidelines for interface design, a toolkit for the low effort implementation of the user interface, and a programming library for the inclusion of speech and Braille in applications. A case study of the development of one such dual interface application is also presented.

1 Introduction

Graphical user interfaces were designed to increase the usability and to improve the functionality of applications [1]. However, a graphical interface is obviously a barrier for blind and visually-impaired people, although assistive devices (primarily speech synthesis, Braille displays, Braille and Qwerty keyboards, voice recognition and screen readers [2,3]) solve many of the problems they come up against. Nevertheless, the use of assistive technologies does not in itself guarantee that blind users will be able to access the application. To this end, the applications need to be built following special design rules [4,5].

Hence the adequacy of a graphical user interface for blind people involves some specific usability requirements for these users. This means that (i) the designer has to present the interface information in an appropriate structure that can be understood by blind users, (ii) none of the meaning of the interface content should be lost during the adaptation and (iii) the information supplied should be adapted for the assistive devices used by blind people. These requirements have an impact on HCI modeling. HCI modeling involves the creation of the task, domain, dialog, presentation, platform and user interface models [6], the design of which is therefore affected by this adaptation.

Based on the above, this paper describes a framework for developing graphical user interfaces for blind people based on usability requirements acquired from our experience in developing this kind of applications [7,8,9].

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 1031 – 1038, 2006.

The paper is organized as follows. Section 2 introduces related work. Section 3 presents usability requirements for blind people. Section 4 describes a framework for blind user interfacing. Section 5 presents an application developed using this framework and the paper ends with some general conclusions.

2 Related Work

The adaptation of an application designed for sighted users to the needs of blind users has the drawback of the original dialog design addressing the specific needs and abilities of sighted users. Also the syntax has to be restructured to meet the needs and abilities of blind people, and this is not always possible [10]. A clear example of this would be a musical score editing application. For a blind user to be able to use such an application, the blind user-adapted tasks and dialog process need to be defined.

A better solution is to create applications with special-purpose interfaces for blind users or with dual interfaces. Toolkits like HAWK [11], which provides a set of standard non-visual interaction objects and interaction techniques that have been especially designed to support high quality non-visual interaction, have been created to facilitate this process. Another tool aiding the development of dual interfaces suitable for both sighted people and blind users is the User Interface Management System (UIMS) [10]. HOMER [12] is an example of an UIMS that improves the development of dual user interfaces.

In any case, we have found, as pointed out by [13], that in order to design an interface suited for use by blind people it is essential to examine the fundamental accessibility issues for such users and define appropriate usability guidelines for them. Not only should these guidelines be formulated as general design principles or low-level and platform-specific recommendations, but they should also be based on experimental evidence [14].

3 Usability Requirements for Blind People

When specifying usability requirements for blind people, it is important to bear in mind that these requirements have to coexist with those of other user types in accordance with Design for All principles (equitable use, flexibility in use, simple and intuitive use, perceptible information, error tolerance, low physical effort, size and space for approach and use) [15] and standard HCI dialog principles (suitability for the task, self-descriptiveness, conformity with user expectations, suitability for learning, controllability, error tolerance, suitability for individualization) [16].

The usability requirements for blind and non-blind people differ to a large extent. Firstly, there is the fact that blind people cannot perform all tasks at the current state of technology. For instance, a blind person cannot as yet drive a car, which means that this task would not have to be adapted for this group of people. This question is referred to as *task adequacy*.

Secondly, blind people and non-blind people use different dimensional access schemes. For sighted users, it is a 2+1 scheme: the user interface objects are distributed in two-dimensional regions (the screen, a window, etc.) and their position typically provides additional semantics about each object. The "+1" component represents the transition from one 2D region to another (navigation), which is performed by user actions. The dimensional access scheme of blind people is 1+1. Devices for blind people (speech and Braille displays) are one-dimensional, which means that user interface objects are presented using a list structure and that the navigation is towards another list. This is what we have called the *dimensional trade-off*.

Thirdly, there is the requirement of maintaining a *behavior equivalence*. A non-blind person interacts using direct manipulation of screen objects, which is not possible for a blind user using assistive technologies. Thus, the interaction process should define how to make each interactive object accessible for the blind. Extra actions need to be added for all objects, including the messages that must be sent to the assistive technologies, the mechanism for accessing each part of the item, and any browse functions that can be performed.

Fourthly, it is not enough just to adapt each object for blind users. The information provided to blind users during task performance also has to be adapted, assuring that there will be no semantic loss during the process, because relevant information may be conveyed by means of images or the spatial distribution of objects located in a window. This is what we have termed *semantic loss avoidance*.

Finally, the assistive technologies used by blind people are not standardized. Speech devices differ both in the functionality they offer and the programming interfaces that have to be used, and the same applies to Braille displays. Applications must deal with this diversity of devices. We call this the *device independency requirement*.

4 A Framework for Blind User Interfacing

The above-mentioned requirements have to be dealt with when applying the modeling approach of user interface development [6], which is based on the task, domain, dialog, presentation, platform and user models.

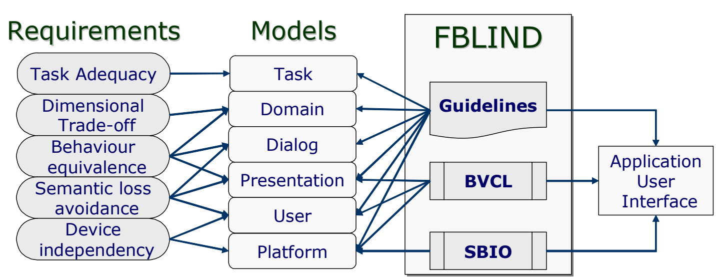


Fig. 1. Blind user requirements, HCI models and FBLIND

Based on the experience acquired over the years, we have developed a framework for the development of dual user interfaces for sighted and blind people: FBLIND (Framework for BLind user INterface Development). Fig. 1 shows the relationship between the blind-user requirements, the HCI models and our framework.

FBLIND has three main components: a set of user interface design guidelines, an interface development toolkit consisting of automatically adapted user interface objects (BVCL) and a programming library providing support for speech and Braille input and output (SBIO).

4.1 Blind User Interfacing Guidelines

These guidelines are based on the recommendations of the Spanish and international standards on software accessibility [4,5]. These recommendations, in whose development we are participating, along with other tips that we have picked up over the years have been incorporated into the proposed framework.

Our guidelines affect all the HCI models to a greater or lesser extent and contain useful information for all the stages of interface design and implementation. One very important aspect of these guidelines is that the application should be compatible with screen readers and other assistive technologies that blind people use to access graphical user interfaces. Below is an overview of the main issues for each of the models.

In the case of the *task model*, there is only one constraint: the tasks described should be checked for incompatibility with the capabilities of blind people using assistive technologies. This caters for the task adequacy problem.

The *domain model* is affected by the dimensional trade-off and the behavior equivalence requirements. To solve the dimensional trade-off, the sequence of windows and their content need to be defined hierarchically as a series of trees linked by transitions between windows to be compatible with 1+1 navigation. As regards behavior equivalence, the designer should define the information about what happens when transitions take place between tree levels or trees to be presented to the blind user as speech or Braille.

For the *dialog model*, the actions to be taken are determined by the behavior equivalence and semantic loss avoidance requirements. To solve the behavior equivalence problem, the dialog model should take into account that the main input device is the keyboard (or equivalent), and it should define each user interface object's interaction process and the access mechanisms to these objects' properties. With respect to semantic loss avoidance, a number of additional interaction commands should be incorporated to improve user operations, such as informing users about where they are in the application, repeating the last message, spelling the last message, etc.

The *presentation model* is subject to the behavior equivalence and the semantic loss avoidance requirements. With respect to the behavior equivalence, the speech and Braille presentation of each of the application's standard and special-purpose user interface objects should be defined. In terms of semantic loss avoidance, several detail levels will first have to be established for this information so that the interface is suited for both novice and experienced users. Additionally, special attention should be paid to non-standard objects and objects with a high graphical content to assure that no semantic information is lost.

The *platform model* is affected by the device independency requirement. Dual interface applications should provide speech output, Braille output or, preferably, both types of output. The platform should provide services for this functionality, and, in order to deal with device independency, it should provide standardized API for speech and Braille devices. The platform model, then, should contain information about the capabilities of those APIs.

Finally, the *user model* has to deal with the semantic loss avoidance and the device independency requirements. In particular, the user model should represent configuration parameters for the functionality required to solve both these problems, such as speech parameters (speed, tone, volume…), Braille parameters (cursor type, Braille code type…) and help level (for novice, intermediate or advanced user).

4.2 BVCL

The usability requirements for blind people applicable to the presentation, platform and user models have been built into a toolkit, called BVCL (Blind user adapted VCL), which is an extension of the Borland C++ Builder development environment and has been developed as a result of three Master's degree dissertations.

BVCL incorporates user interface objects with automated presentation in speech and Braille (presentation model), which saves development time. It also eases the programming of additional speech and Braille messages (platform model).

Finally, the specific messages generated in speech and Braille for each user interface object can be easily configured through BVCL, and it also incorporates layout managers to automate the layout and sizing of the user interface objects according to user preferences on font sizes, colors, etc. (user model).

4.3 SBIO

SBIO (Speech and Braille Input / Output) consists of a series of programming components that improve speech and Braille input and output. It is a package for Microsoft Windows, as this is one of the platforms most commonly used by blind people. It was developed jointly by Microsoft, the ONCE (the Spanish National Organization for the Blind), and the UPM (Madrid Technical University).

SBIO contains four components: the Microsoft speech API, a Braille API, a set of ActiveX components that offer a simplified and unified speech and Braille interface (SBIO-Core), and a dynamic link library (SBIO-Lib) that facilitates the use of IOSB.

5 An Environment for Bilingual Dictionaries with a Dual Interface

DABIN is an environment that includes bilingual dictionaries with a generic interface (Fig. 2) combining visual and blind interfaces (using speech synthesis and Braille). Its basic function is to search for words in one language and get their translation in another language from a range of bilingual dictionaries. Users can choose the source and target languages of the translation, as well as the user interface language.

All the tasks in *task model* were checked to ascertain whether they were adequate for blind users and accessible by means of the keyboard, speech synthesis and Braille display, and that they did not require hand-eye coordination or simultaneous control of different visual items.

In the *domain model*, DABIN has a hierarchical structure based on linked trees that can be easily explored in both the 2+1 and 1+1 access schemes. It should be noted that the tree structure for each software window is simple and quite shallow. Because blind users must necessarily navigate by levels within the tree and cannot "take a look" to find out where to go, the software windows are designed to assure that their navigation tree is not overly complicated.

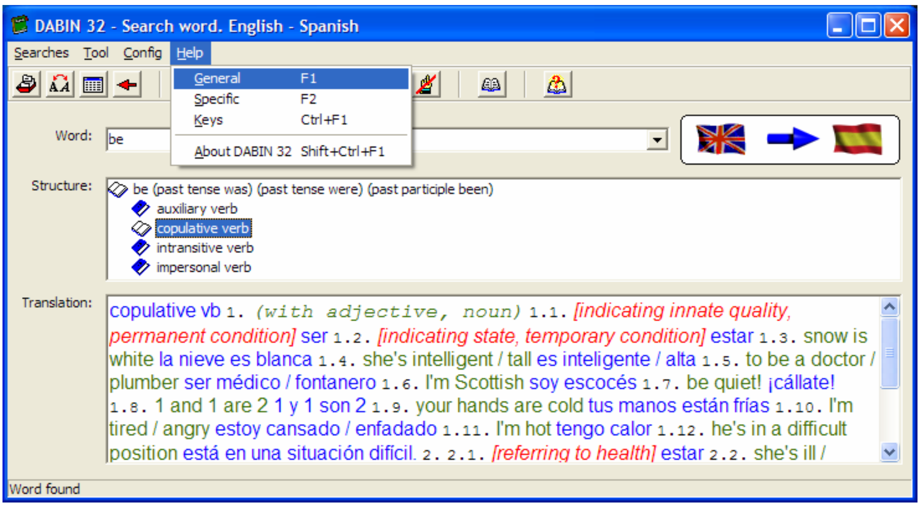


Fig. 2. DABIN's graphical user interface

In the *dialog model*, mechanisms were enabled to allow blind users to do all the tasks using key combinations or system menus. The interaction procedure for each user interface object was also defined for adaptation to blind people's needs. In addition to the standard user interface objects included in BVCL, a new complex object was needed: a translation browser, used to display the translation (bottom box of Fig. 2). We defined an interaction procedure for this component to make the reading of a translation a natural activity for blind users. This procedure is based on enabling several ways of getting the information using the keyboard only, allowing users to move forwards and backwards over the translation, select text and start a search for the word which the cursor is over. Behavior compatible with standard edit boxes was applied in the translation browser design.

Commands to access additional semantic information, such as user location, repetition of the last message, spelling or context-sensitive help, have been added to the system interaction techniques to prevent semantic information loss.

To build a dual interface the *presentation model* considered the visual appearance (following the Windows standards) and the non-visual appearance (focusing on sound –speech synthesis– and haptic –Braille– presentation). The sound and haptic interfaces for each user interface object were defined using BVCL and SBIO. The output in response to the different user actions on the translation browser was also defined in detail.

Finally, we identified three user types in the *user model*: sighted users (no special-purpose adaptation required), visually impaired users (the colors and the sizes of the texts can be configured), and blind users (the voice parameters can be configured for each of the languages as can the general appearance of the Braille).

To check that the information provided by the system was correct, several users of the three identified types tested this interface. As a result, the model was refined until the assistive technologies provided equivalent semantic information to what was displayed on screen.

6 Conclusions

We believe dual interfaces development to be the best option for efficiently developing and building user interfaces that satisfy the needs of both sighted and blind users. To this end, we have defined specific requirements that take into account all the problems that blind users face and that coexist with the standard dialog principles and Design for All guidelines. We have built these requirements into the HCI models. We have also presented a framework (FBLIND) to improve the development of these user interfaces. This framework includes a set of *guidelines*, which help the developer to focus on relevant issues during the construction of the six HCI models; a *toolkit* for auto-adapted user interface objects, which greatly reduces the implementation costs of dual user interfaces, and an *API* for the use of assistive devices, which provides an easy way to implement speech and Braille output. The use of this framework has been illustrated by means of DABIN, a software environment for bilingual dictionaries for use by blind and sighted users.

More tool support should be provided for tasks that are not yet covered by the toolkit and the API. A CASE tool covering the whole user interface development life cycle, providing support for manual tasks would be the ideal thing.

References

1. Cox, K., Walker, D.: User interface design. Prentice Hall (1993)
2. Baum Engineering SRL: Welcome to Gnopernicus!, available at URL: `http://www.baum.ro/gnopernicus.html` (consulted January 10, 2006)
3. Freedom Scientific BLV Group: JAWS(r) for Windows(r) Quick Start Guide, 11800 31st Court North, St. Petersburg, Florida 33716-1805 USA (2005)
4. ISO: Ergonomics of human-system interaction – Guidance on accessibility for human-computer interfaces, Technical Specification ISO TS 16071 (2003)
5. AENOR: Computer applications for people with disabilities. Computer accessibility requirements. Software (Spanish title: Aplicaciones informáticas para personas con discapacidad. Requisitos de Accesibilidad al Ordenador. Software), Spanish Standard UNE 139802 (2003)

6. Szekely, P., Luo, P., Neches, R.: Beyond Interface Builders: Model-Based Interface Tools. In: Proc. of InterCHI'93, ACM Press, New York (1993) 383-390
7. Alonso, F., Frutos, S., Montes, C.: Interfacing for the Blind. The DILE Model. In: Proc. CSUN´s Twelfth Annual International Conference on Technology and Persons with Disabilities, Los Angeles, March, (1997)
8. Alonso, F., Frutos, S., Fuertes, J. L., Martínez, L. A., Montes, C., Sánchez, A.: Un Diccionario Automático Bilingüe para Invidentes (DABIN). Integración 25 (1997) 14-31
9. Alonso, F., Barreiro, J. M., Caraça-Valente, J. P., López, G., Montes, C., Navajo, R. J.: An Electronic Newspaper for the Blind. In: Proc. XV IFIP World Computer Congress, Computers and Assistive Technology ICCHP '98, Vienna (Austria) and Budapest (Hungary) (1998)
10. Savidis, A., Stephanidis, C.: Developing dual user interfaces for integrating blind and sighted users: the HOMER UIMS. In: Proceedings of the SIGCHI conference on Human factors in computing systems, Denver, Colorado, United States (1995) 106-113
11. Stephanidis, C., Savidis, A.: Interface Development Toolkits for Non-visual and Switch-based Interaction. ERCIM News, 46 (2001), available at URL: `http://www.ercim.org/publication/Ercim_News/enw46/stephanidis2.html`.
12. Savidis, A., Stephanidis, C.: The HOMER UIMS for dual user interface development: Fusing visual and non-visual interactions, Interacting with Computers 11(1998) 173-209
13. Morley, S.: Design and Evaluation of Non-Visual Information Systems for Blind Users, PhD, University of Hertfordshire, UK (1999)
14. Casali, S. P.: A physical skills-based strategy for choosing an appropriate interface method. In: Edwards, A. D. N. (ed.): Extra-ordinary Human Computer Interaction: Interfaces for people with disabilities, Cambridge University Press (1995) 315-342
15. Connell, B. R., Jones, M., Mace, R., Mueller, J., Mullick, A., Ostroff, E., Sanford, J., Steinfeld, E., Story, M., Vanderheiden, G.: The Principles Of Universal Design, NC State University, The Center for Universal Design (1997)
16. ISO: Standard 9241-10:1996, Ergonomic requirements for office work with visual display terminals (VDTs) – Part 10: Dialog principles, International Organization for Standardization (1996)

An Approach for Direct Manipulation by Tactile Modality for Blind Computer Users: Development of the Second Trial Production

Shigenobu Shimada[1], Masami Shinohara[2], Yutaka Shimizu[3], and Makoto Shimojo[4]

[1] Tokyo Metropolitan Industrial Technology Research Institute,
Research and Development Department, Nishigaoka 3 13 10, 1158586 Tokyo, Japan
shimada.shigenobu@iri-tokyo.jp
[2] National Institute of Advanced Industrial Science and Technology,
Institute for Human Science and Biomedical Engineering, 3058568 Tsukuba, Japan
m-shinohara@aist.go.jp
[3] The University of Electro–Communications, Department of Systems Engineering, Chofugaoka 1 5 1, 1828585 Tokyo, Japan
shimizu@sys.uec.ac.jp
[4] The University of Electro–Communications, Department of Mechanical Engineering and Intelligent Systems, Chofugaoka 1 5 1, 1828585 Tokyo, Japan
shimojo@mce.uec.ac.jp
http://www.rm.mce.uec.ac.jp

Abstract. A basic device combining a tactile display function and a touch position sensing function is proposed. The trial device consists of two major components, a tactile graphic display and a six–axis force/torque sensor. The force sensor measures six dynamic values generated by touch action on the display surface and a PC estimates the touch position based on the data. Since the defects of the first trial production are the weakness of the touch surface, an assembly error, and the measurement error of six–axis force/torque sensor, they are solved in the second trial production. The effect of a contact force on the estimated position are examined respectively by a vertical component and a horizontal component. It is shown from the above experimental results that the second trial production is practically sufficient estimated position accuracy.

1 Introduction

Most of the tactile display devices developed in the past were only unidirectional communication tools, that is only display devices [1,2,3]. If a bidirectional communication function could be provided, graphic information would become more accessible for visually impaired personal computer (PC) users and stimulate their creative activity. So, we have developed a tactile input/output device which combined the tactile display and a six–axes force/torque sensor [4,5]. The six–axes force/torque sensor measures the force which the user applies to the tactile display, and PC estimates the touch position on the tactile display which is applying force based on the measurement value. The click function to specify the arbitrary coordinates on the tactile display by an empty hand is realized. Furthermore, scrolling of the screen is enabled from measurement of force and it become possible to search all the areas of the PC screen on the tactile display with 48×32 dots. Based on these, the inspection of the file operation by GUI

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 1039–1046, 2006.

(Graphical User Interface) operation or a web page will be attained. However, in the first trial production, when a horizontal component of force is added, the estimated position and the touch position is shift greatly. As a result of performing analysis and examination, the cause of the position shifts are able to consider the weakness of the touch surface, the assembly error, and the measurement error of the six–axis force/torque sensor, etc. In this paper, the new device which solved the above mentioned three problems is developed. The correction method of the touch position is newly drawn using the knowledge acquired with the first trial production with development of a second trial production. We investigate the accuracy of detecting method when the touch action is placed on the second trial production.

2 Second Trial Production

There were the following three problems in the first trial production.

1. Although the touch surface of the first trial production is constituted by the top plate of tactile pin module, the module is a product made of resin. Therefore, the touch surface will be distorted when the touch surface is weighted.
2. The assembly error exists in our device, and the device has influence which cannot be disregarded for the estimated accuracy.
3. Used six–axes force/torque sensor includes the 5% error in a measurement value when the sensor is weighted between multi axes.

In order to solve the item 1., the tactile pin module is put into the metal box. 1536 holes are prepared in the top plate of the metal box, and Teflon processing is given to the surface. In order to solve the item 2., the correction formula is newly drawn. It explains in full detail behind. The item 3. is necessary on the structure of the sensor. Therefore, although influence of the item 3. is not completely eliminable, we changed the rating of six–axes force/torque sensor and the error is reduced relatively. Below, the tactile display and six–axes force/torque sensor are described.

2.1 Tactile Graphic Display

Four bimorph modules (SC–10, provided by KGS Inc.) are used to assemble the tactile graphic display. Each module has a 32×12 arrangement of pins with 2.4 mm spacing. The pin is 1.3 mm in diameter and must be pushed down with the force of at least 0.098 N in order to touch the surface. The total tactile surface has the 32×48 arrangement activated at a time. The height of the tactile pin when activated by a data signal is 0.7 mm. The refresh time for the entire surface is about 50 ms. The device interfaces with the host PC through a USB or serial communication. The top right of Fig. 1 shows the component of the tactile display.

2.2 Force/Torque Sensor

The six–axes force/torque sensor is used in the experiments. The sensor measures the force and the torque generated by touch action. The touch action contains important information such as touch position, push direction and push strength. These quantities

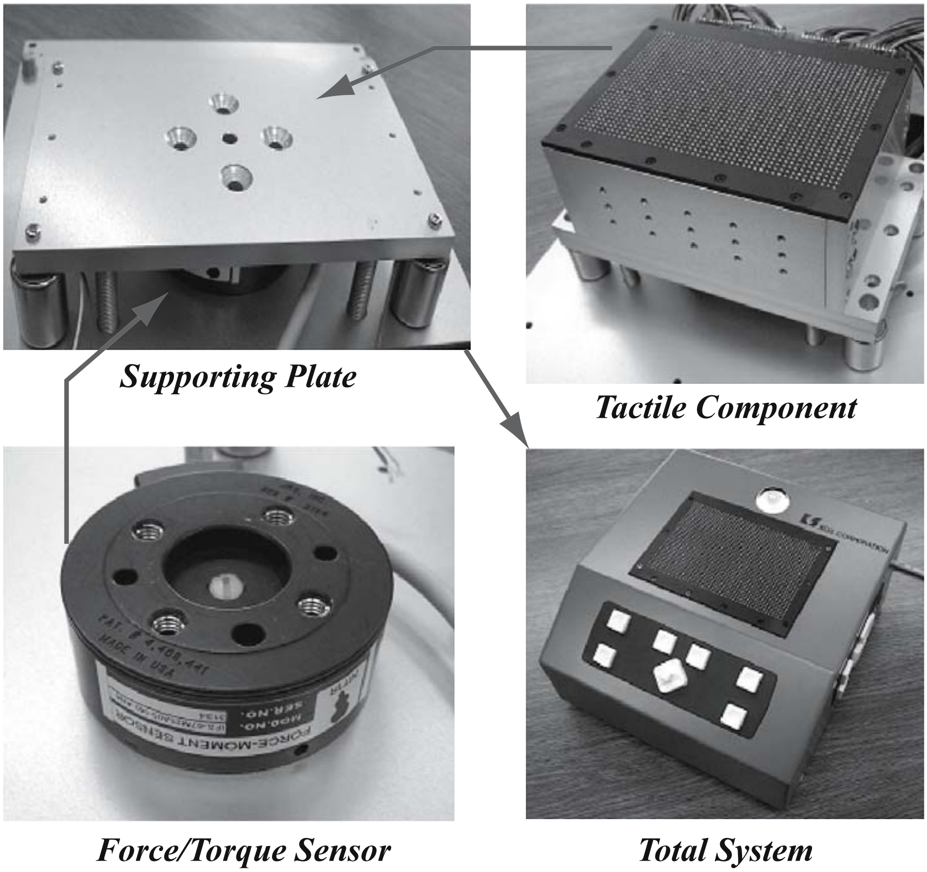


Fig. 1. Actual structure of the second trial production is shown. A display component is fixed on a solid plate firmly connected to a six–axis force/torque sensor.

are estimated based on the measured data and used for discriminating proper touch actions and designing HCI (Human Computer Interaction) functions such as mouse like function, touch scroll function and so on. The sensing ranges are respectively 25 N (in regard to the x– and y–axis), 50 N (in regard to the z–axis) in force, and 1 Nm (in regard to the x– and y–axis), 0.5 Nm (in regard to the z–axis) in torque. The display is fixed on a solid plate firmly connected to the force/torque sensor as shown in Fig. 1. The sensor is set roughly below the center of the display component.

3 Principle of Touch Position Estimation

3.1 Basic Equation

Fig. 2 helps to explain the operational principle adopted here. When the user pushes the surface Π with force $\boldsymbol{f}$ at $\boldsymbol{r}$, the following simultaneous equations are formed:

$$\boldsymbol{m} = \boldsymbol{r} \otimes \boldsymbol{f}, \tag{1}$$

$$(\boldsymbol{h}, \boldsymbol{r} - \boldsymbol{h}) = 0 \tag{2}$$

where Eq. (1) is force–moment equation, Eq. (2) define the touch surface, $\boldsymbol{r}$: touch position vector, $\boldsymbol{f}$: force vector, $\boldsymbol{m}$: force moment vector, $\boldsymbol{h}$: normal vector defining the touch surface, $\otimes$: vector product, and $(,)$: scalar product. The touch position is theoretically obtained as the solution of Eq. (1) and Eq. (2):

$$\boldsymbol{r} = \frac{\{\boldsymbol{h} \otimes \boldsymbol{m} + (\boldsymbol{h}, \boldsymbol{h})\boldsymbol{f}\}}{(\boldsymbol{h}, \boldsymbol{f})}. \tag{3}$$

Since $\boldsymbol{r}$ is a three–dimensional solution, it is necessary to transform into the touch surface coordinate system. If a transformation matrix is placed with $\boldsymbol{A}$ and a parallel translation vector is placed with $\boldsymbol{c}$, the transformation result are formed:

$$\boldsymbol{r}' = \boldsymbol{A}\boldsymbol{r} + \boldsymbol{c}, \tag{4}$$

where

$$\boldsymbol{A}^t\boldsymbol{A} = \boldsymbol{I}, \tag{5}$$

$$\boldsymbol{A}\boldsymbol{h} = 0. \tag{6}$$

3.2 Estimation of Normal Vector *h*

Normal vector $\boldsymbol{h}$ is estimated from the measurement value. When we set the touch position to $\boldsymbol{r}_i (i \in \boldsymbol{I})$ and set the left side of Eq. (4) corresponding to $\boldsymbol{r}_{ij} (i \in \boldsymbol{I} \& j \in \boldsymbol{N}_i)$, let the vector $\boldsymbol{h}$ which minimizes formula (7) equal the normal vector. Under the constraint of Eq. (5), $(\boldsymbol{h}, \boldsymbol{A})$ which minimizes the following equation (7) leads to the optimal estimation in the meaning of least squares estimation.

$$Q = \sum_i \sum_j (\boldsymbol{r}_i - \boldsymbol{A}\boldsymbol{r}_{ij} - \boldsymbol{c}, \boldsymbol{r}_i - \boldsymbol{A}\boldsymbol{r}_{ij} - \boldsymbol{c}) \tag{7}$$

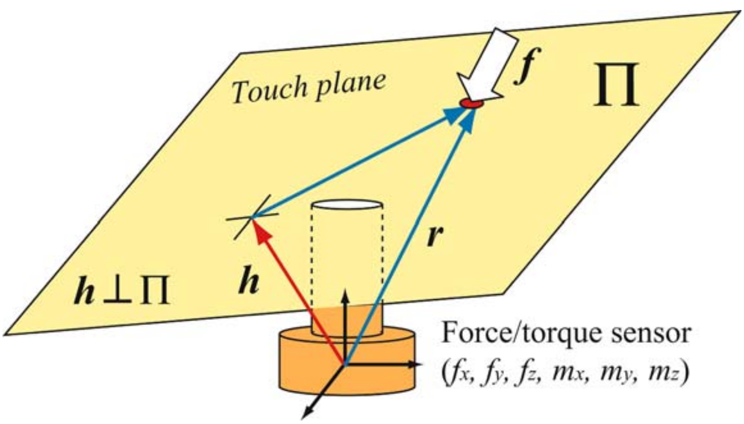


Fig. 2. Touch position sensing principle

By the way, although $\boldsymbol{h}$ changes delicately with the bolting torque of an assembly etc., it is thought that it is usually fixed. Moreover, when distortion is large, it is also considered that the estimated position cannot be corrected in the affine transformation of Eq. (4). It is desirable to estimate $\boldsymbol{h}$ beforehand from this viewpoint. The estimated $\boldsymbol{h}$ is set to $\boldsymbol{h}$ where the following Eq. (8) is maximized using the above mentioned $\boldsymbol{r}_{ij}(i \in \boldsymbol{I} \& j \in \boldsymbol{N}_i)$.

$$R = 1 - \frac{\sum_i \sum_j (\boldsymbol{r}_{ij} - \overline{\boldsymbol{r}}_i, \boldsymbol{r}_{ij} - \overline{\boldsymbol{r}}_i)}{\sum_i \sum_j (\boldsymbol{r}_{ij} - \overline{\boldsymbol{r}}_{ij}, \boldsymbol{r}_{ij} - \overline{\boldsymbol{r}}_{ij})} \tag{8}$$

where

$$\overline{\boldsymbol{r}}_i = Average_j(\boldsymbol{r}_{ij}), \overline{\boldsymbol{r}}_{ij} = Average_{i,j}(\boldsymbol{r}_{ij}). \tag{9}$$

However, $\boldsymbol{h}$ is obtained within the area which is predicted from structure.

3.3 Estimation of A and c

$\boldsymbol{r}_{ij}$ is calculated by substituting the above mentioned $\boldsymbol{h}$ into Eq. (3). $\boldsymbol{Q}$ is obtained by substituting the the obtained $\boldsymbol{r}_{ij}$ into Eq. (7), and we obtain $\boldsymbol{A}$ and $\boldsymbol{c}$ which minimize Eq. (10).

$$J = Q + (\lambda_1, \lambda_2)\boldsymbol{A}\boldsymbol{h} \tag{10}$$

The right hand second term is the constraint of Eq. (6). The reason for not adding Eq. (5) to the constraint is that the output of six–axes sensor includes the 5% error in the measurement, when multi–axial tension is applied. If Eq. (5) is added to the constraint, in order not to allow even small elasticity, we cannot expect the improvement in the estimated accuracy.

4 Experimental Investigation

The effect of the touch force on the estimated position are examined by a vertical component of force and a horizontal component of force, respectively. These characteristics are the foundations of accuracy evaluation our device.

4.1 Experiment 1: Estimation Accuracy Using Vertical Component of Force

We first examine the accuracy of the estimated position on the touch surface using an algorithm based on the principle mentioned above. As experimental conditions, the touch plane is delimited to the lattice and a vertical component of force is added with twenty five measurement points. In addition, we use the six kinds of weights as a load, because it is thought that the user operate our device according to not the same force but various force in practical use. Therefore, six estimated positions are obtained from six vertical component of forces in one measurement point. First, the touch plane is assumed to be orthogonal to $\boldsymbol{h}$, and the distortion correction is not done. On these conditions, Fig. 3 shows the estimated positions (denoted by square) in comparison with the touch positions (denoted by circle). The estimated position is estimated in the direction of minus

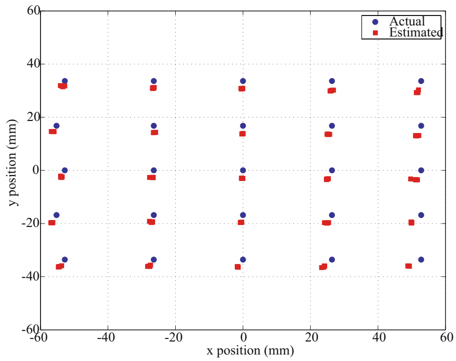


Fig. 3. The estimated positions (denoted by square) in comparison with the touch positions (denoted by circle) : Conditions (1. the touch plane and $\boldsymbol{h}$ are orthogonal, 2.non correction.)

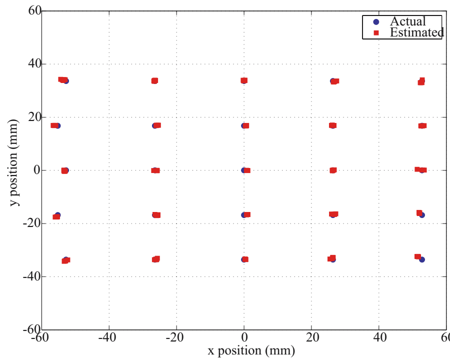


Fig. 4. The estimated positions (denoted by square) in comparison with the touch positions (denoted by circle) : Conditions (1. the touch plane and $\boldsymbol{h}$ are not orthogonal, 2.correction.)

y against the touch position. Moreover, we understand that the influence by the kind of the weight can be disregarded. Second, we assume that the touch plane and $\boldsymbol{h}$ are not orthogonal, we apply the correction of $\boldsymbol{h}$ and the affine transformation. Fig. 4 shows the experimental result. The normal vector $\boldsymbol{h}$ is calculated with $\boldsymbol{h} = (-0.8, -1.2, 83.8)$, that is, the position difference to the direction of minus y is reflected. By performing the transformation, all the estimated position can be agreed with the touch position.

4.2 Experiment 2: Estimation Accuracy Using Horizontal Component

Fig. 5 shows the experimental apparatus. The horizontal component of force is applied by a tilted weight. The experimental conditions are follows:

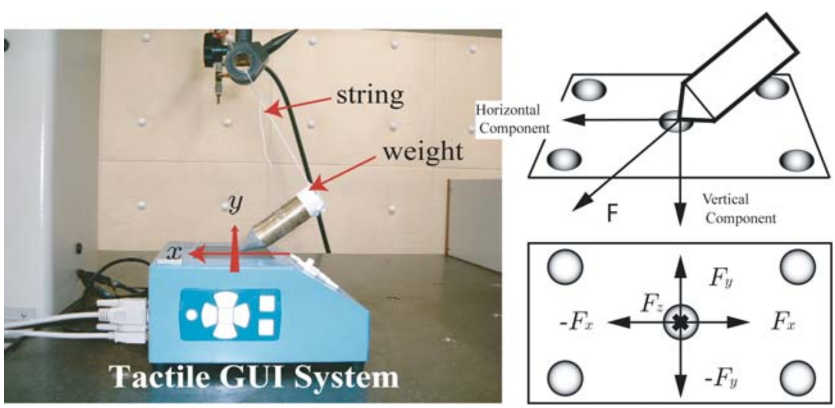


Fig. 5. Experimental apparatus using horizontal component of force

Table 1. Experimental result: The evaluation object is a difference between actual touch position and estimated one. A unit is mm.

	Non Correction					Correction				
	Point 1	Point 2	Point 3	Point 4	Point 5	Point 1	Point 2	Point 3	Point 4	Point 5
$-f_x$	1.851	1.737	1.028	1.758	1.447	2.74	0.866	1.083	1.101	1.030
$-f_y$	2.47	2.36	1.281	0.423	2.39	2.74	1.055	1.640	0.478	0.563
$+f_x$	2.03	2.81	1.257	0.457	4.12	1.422	1.238	0.973	1.014	2.18
$+f_y$	1.374	2.32	0.965	1.613	2.86	1.488	0.753	0.298	1.370	1.546

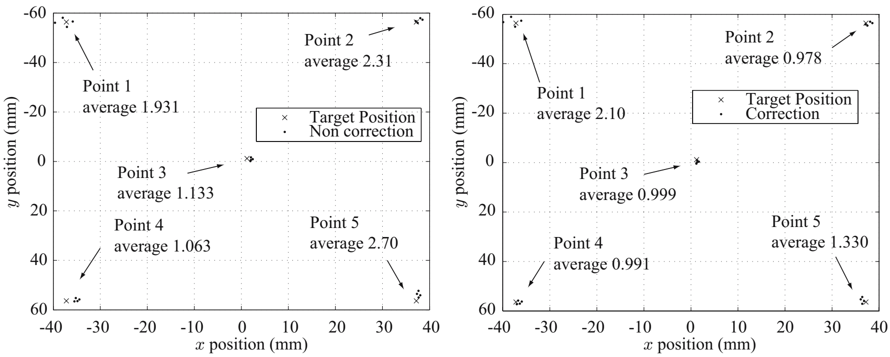


Fig. 6. Touch position data using the horizontal component of force

1. Weight is 500 g, and its shape is a column of 82 mm and a cone of 29 mm.
2. Forces are measured by the six–axes force/torque sensor.
3. Measurement carried out at five points, i.e. top left (Point 1), top right (Point 2), center (Point 3), bottom left (Point 4), bottom right (Point 5).
4. Each measurement point is measured in four horizontal directions.

The experimental results are shown in Fig. 6. The evaluation object is a difference between actual touch position and estimated one. Table 1 shows the evaluation objects in each measurement points. The average of all measurement points of the non correction is 1.828 mm, and the average of all the measurement points of the correction is 1.279 mm. The accuracy of estimated position has been increased 30% in comparison with non correction.

5 Conclusion

There is a tactile display as one of the transmission of information apparatus for visually impaired persons. However, the function of display is almost limited only one–direction. In this study, a basic device combining the tactile display function and the sensing function was proposed. We also proposed the new device which solved the weakness of the touch surface, assembly error, and the measurement error of the six–axis force/torque sensor. In the second trial production, the new correction method is drawn, and the experiment of position estimation is conducted. As the result of experiment, the improvement of the estimated position accuracy is admitted compared with the first trial production. In the future, we pursue the possibility of our device by applying to applications of the cooperation of voice information and the fusion of tactile map, etc.

References

1. Gardner, J. A., Bulatov, V., "COMPLETE ACCESS TO ALL PAPER AND COMPUTER FORMS, TABLES, AND CHARTS," in Proc. of the 2004 CSUN International Conference on Technology and Persons with Disabilities, pp. 17–20, 2004
2. Shinohara, M., Shimizu, Y., Mochizuki, A., "Three–dimensional tactile display for the blind," IEEE Trans. On Rehabilitation Engineering, Vol. 6, No. 3, pp. 249–256, 1998
3. Wagner, C. R., Leederman, S. J., Howe, R. D., "A tactile shape display using RC servomotors," in Proc. of Satellite Meeting of the Annual IEEE VR'02, pp. 354–356, 2002
4. Shimojo, M., Shinohara, M., Tanii, M., Shimizu, Y., "An approach for direct manipulation by tactile modality for blind computer users: Principle and practice of detecting information generated by touch action," Computers Helping People with Special Needs (ICCHP2004), Springer–Verlag, pp. 753–760, 2004
5. Shimada, S., Shinohara, M., Shimizu, Y., Shimojo, M., "Regarding Accuracy of tactile input/output device," in Proc. of International Conference on Sensor Technologies (ICST2005), pp. 438–443, 2005

Making Nonaccessible Applications Accessible for Visually Impaired

Tomáš Zahradnický and Róbert Lórencz

The Czech Technical University,
Karlovo Náměstí 13, 12135 Prague, Czech Republic
zahradt@fel.cvut.cz, lorencz@fel.cvut.cz
http://cs.felk.cvut.cz/~zahradt/

Abstract. Apple Human Interface Guidelines state that all applications should be accessible in order to provide the best user experience. Not all software vendors honor this statement and if an application does not support accessibility, people with disability or a special need can hardly use it or cannot use it at all. The purpose of this paper is to present a method that is capable of extending an already existing applications to support accessibility without the necessity to have access to the source code of the application. Paper proposes the use of Mach code injection techniques to load access enabling code into the context of the access enabled application and describes ways that are used to make the application accessible. There is also a case study of proposals offered by this paper on a portion of ProTools, a professional audio editing software that makes its mix table window accessible.

1 Introduction

Accessible applications provide useful information about their user interface and its state information to an assistive application in a defined way and the assistive application presents that information to the user. User decides what to do and instructs the assistive application that asks in turn accessed application to perform the desired action. The process of access enabling (accessorizing) of an application provides alternative to mouse-driven user interface and offers support for full keyboard navigation. Accessible applications can be used by millions of people with disability or a special need and even people who do not necessarily have any disability can benefit from it.

Accessorizing is a task that takes part in application development cycle and is very easy for applications that use only the user interface elements (widgets) provided by the operating system. If the application defines its own widgets, it is required by Apple Human Interface Guidelines [4] to provide necessary accessibility support. Whether to respect or not this document is always a decision made by the application developer and if the developer decides to not support accessibility for his custom views, visually impaired people are not able to use his product. This paper presents an approach that could be taken to make an already existing nonaccessible application accessible with the help of Mach code injection.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 1047–1054, 2006.

2 State of the Art

In Mac OS X 10.2, accessibility support got greatly enhanced and Accessibility framework has been introduced. This framework provides accessibility protocol Carbon and Cocoa applications adopt and also provides an API that is used by assistive applications to query and direct user interface in another running application. Since then, each system widget has been extended to conform to the accessibility protocol [5] to provide defined response to the assistive application. With the announcement of Mac OS X 10.4, VoiceOver [5] technology has been introduced to provide a fully spoken user interface and visually impaired users are able to navigate through the user interface seamlessly. As long as the user interface is made of standard Carbon, Cocoa, or Java either AWT or Swing widgets, it is automatically accessible and only few more work to add objects' descriptions and to bind title user interface elements with their corresponding widgets is expected from the developer.

On the other hand, custom widgets may be sometimes necessary and such widgets have to adopt the accessibility protocol to ensure that they properly respond to queries and requests from the assistive application. The way how they do so depends on the framework used and it is wrapped in the following table:

Table 1. Accessibility in Mac OS X frameworks

Framework	Implementing accessibility via
Carbon	`kEventClassAccessibility` Carbon events
Cocoa	`NSAccessibility` protocol
Java	`javax.accessibility` package

Each accessible widget has an associated accessibility object (`AXUIElement`) that stores information about widget and is used during the communication with the assistive application and accessibility API. Accessibility objects are organized in hierarchies and one object can embed one or more accessibility objects. Such hierarchy of objects allows an application to contain its windows accessibility objects, menu bar accessibility object, and some additional objects. Contents of any accessibility object that is associated with an onscreen widget can be comfortably monitored with Accessibility Inspector application that is installed together with Mac OS X Developer Tools. Accessibility Inspector is also used to debug the `AXUIElement` contents and hierarchy eg. during the pruning process.

Each accessibility object has attributes such as its role, title, description or its parent. Besides attributes, accessibility objects can perform actions that are also presented to the assistive application and user can ask accessibility object to perform such action. A list of mandatory attributes and actions for object's roles is described in [5].

3 Access Enabling New Applications

The process of access enabling of applications is not a very difficult task and consists mostly of adding descriptions to already existing widgets and binding standalone text titles with widgets descriptions semantically belong to. All of the typical accessorizing scenarios, that are for Carbon, Cocoa and Java applications are described in the following three subsections.

3.1 Accessorizing Carbon Applications

Most Carbon applications are accessible if they use standard widgets but there is still need to supply widget descriptions and possibly their title elements. Accessibility in Carbon applications is attained in the five following steps [9]:

1. *Preparing* — in this stage, keyboard navigation is designed, help tags are provided and window proxies are established so they can connect a window to a file. Having a window proxy set automatically enables accessibility for document windows.
2. *Context related accessorizing* — provides auxiliary attributes and also binds title objects with widgets they logically belong to.
3. *Pruning* — it is necessary to set accessibility as ignored for widgets, that are not meaningful to be accessed such as container views that typically embed other widgets.
4. *View accessorizing* — provides access for all views.
5. *Custom widgets accessorizing* — this is the most complex step that makes custom views adopt the accessibility protocol. This means that a view has to provide its description, role, value, a list of actions and other attributes assistive application may need.

3.2 Accessorizing Cocoa Applications

Cocoa applications are easier to accessorize hence most of the accessorizing work is done at the UI design level and works automatically. Should the need to customize this behavior arises, the following can be used [8]:

1. *UI Design* — title and linked accessibility objects are specified during the Cocoa UI design process, therefore they work automatically.
2. *Setting instance specific attributes* — sometimes it might be necessary to override default settings for title and description and this is typically accomplished by using accessibility APIs.
3. *Subclassing accessorized classes* — subclassing is used to construct custom views and if a view needs to be accessible, it is required to override several Cocoa methods. These methods define whether the object is ignored, provide object's attributes and perform hit testing.
4. *Accessorizing from scratch* — if a completely custom behavior is used, object can override approximately ten accessibility methods and handle all requests from assistive application by itself. That includes providing a list of attributes, a list of actions, hit testing, etc.

3.3 Accessorizing Java Applications

Java applications typically need the Java Accessibility bridge to provide accessible GUI with the help of `javax.accessibility` package. In Mac OS X, this technology is already built into Java 2, Standard Edition and all Java applications that use AWT or Swing windowing models are automatically accessible.

4 Access Enabling Already Existing Applications

This paper presents a method of accessorizing windows in an already existing Carbon (host) application we do not have source code of. We assume that the host uses user interface that does not consist of system widgets and provides its own, non-accessible widgets. In order to provide accessibility in such host, we have to provide the information assistive application requires by means of accessibility protocol. To provide this information, each window has to provide accessibility object hierarchy that can be created by creating `HIViews` that handle all of the necessary events of `kEventClassAccessibility` event class. In further we will call such view an `axView` and its purpose is to provide a placeholder that has no physical appearance but provides information required by the assistive application. Each `axView` responds to the following accessibility events in addition to events requires by `HIView` and its base class `HIObject`:

- `kEventAccessibleGetAllAttributeNames` — this event provides a list of attribute names (such as role, title, description) that will given `AXUIElement` contain. `axView` adds 3 attributes that are:
  - `kAXTitleAttribute` — attribute that holds a object's title.
  - `kAXDescriptionAttribute` — attribute containing object's description.
  - `kAXFocusedAttribute` — attribute that defines whether the object currently has a keyboard focus or not.
- `kEventAccessibleGetNamedAttribute` — the purpose of this event is to provide a value for an attribute specified by its name.
- `kEventAccessibleGetAllActionNames` — event similar to the first one, where the `AXUIElement` is expected to return a list of actions it supports. `axView` provides no default actions but its descendants such as button typically provide `kAXPressAction` action.
- `kEventAccessiblePerformNamedAction` — called when the user has decided to let object perform given action based on its name.
- `kEventAccessibleGetNamedActionDescription` — provides a description for a name based action.

With the help of `axView` and its descendants, we are able to create the missing `AXUIElement` hierarchy in host's windows and therefore provide navigation support. `axView` can serve as a base class for many other views and a typical sibbling monitors the state of the widget we provide access to and updates appropriately — changes its description when the color of the widget changes, for example. An example of a view that provides `kAXPressAction` action is `axButton` subclass that provides an access to a push button beneath it. When the user decides to push this button, `axButton` obtains `kEventAccessiblePerformNamedAction` event for `kAXPressAction` action name. Its event handler resolves this event by synthesizing a click on itself that gets dispatched to the original accessed widget.

4.1 Getting Accessorizing Code Run in Host'S Context

Accessibility notifications can be used to get informed when a new window appears in target application and assuming that we can safely distinguish one host's window from another, we are able to map all custom widgets to `axView` objects and provide access to that portion of the user interface. In order to do that, we need to get accessorizing implementation into host's address space and the easiest way to do so is from a plug-in loaded into the host.

If the plug-in way is not feasible or when the host does not support plug-in architecture, we can build our code as a dynamic library and use function attributes `__attribute__((constructor))` and `__attribute__((destructor))` to get attributed functions run when the library loads and unloads. Dynamic loader (dyld) can be instructed to load specified libraries into each new process when `DYLD_INSERT_LIBRARIES` environment variable is set. Since such behavior might be undesired, this way is neither preferred nor recommended.

Third, and probably the most efficient way depicted at Fig. 1, is to use Mach kernel APIs [1,2] to allocate a block of memory in host, into which we copy a loader code that will be responsible for dynamic library (or bundle) loading and its consequent execution. Such method is possible as long the current process privileges match host's privileges. The code is copied with the Mach virtual memory APIs [2]. Once we have code ready in the host, we have to allocate an another block of memory that will provide a stack space for a new thread we create. Having code copied and stack allocated, we can create a Mach thread in host[1] (1). Math thread is initially in suspended mode so we can set its internal state which features initial register settings such as program counter and stack pointer. Thread's program counter is set to the previously copied code and stack pointer is set as well. Thread's main routine can also get parameters either in registers (for PowerPC based Macs) or on stack (for Intel based Macs) and finally we can start the thread remotely. Mach thread starts off with a creation of a pthread (2) that performs the initialization process and terminates itself.

From our experiences, even some of the "thread safe" marked functions did not like to be called from a Mach thread and therefore the creation of the pthread was necessary. Pthread's purpose is to install a Carbon Event timer that fires at system task level (3), at that it is safe to call any APIs that we need to load bundle into host's context (4), (4') and then exit. Once the installed timer fires, it loads specified bundle (whose full path is one of the arguments passed to the Mach thread) into host's context and calls its main entry point that can do whatever necessary (5).

[1] This method is usually called code injection and is described more in [10,11] under `mach_inject`.

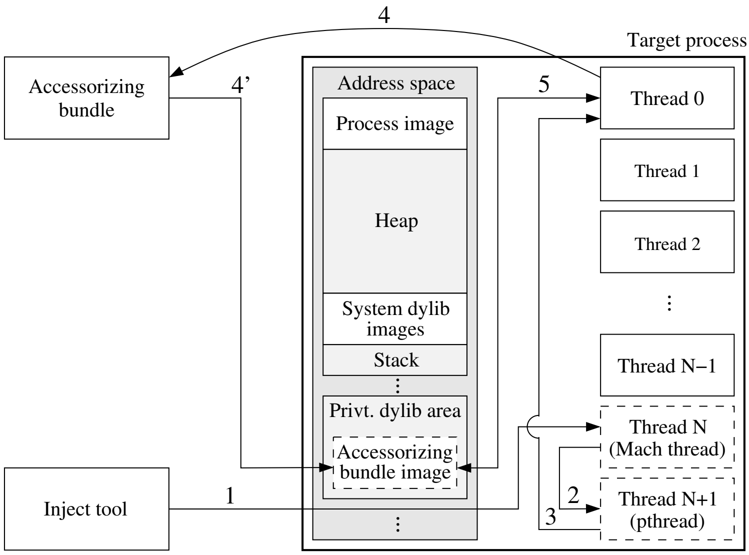


Fig. 1. Loading accessorizing bundle into a process

5 Case Study

The idea presented in this paper was applied on the newest version of ProTools, a professional audio editing software developed by DigiDesign, a division of Avid Technology, Inc. company. ProTools is a good case study sample because its user interface consists mainly of custom widgets that have no relation to the widgets provided by the operating system. Such non-relation is logical since system widgets match neither appearance nor behavior of mix table widgets and ProTools implement these widgets on its own.

The primary focus was to accessorize the contents of ProTools' mixer window that provides software control to a hardware mix table that must be attached to the computer via USB or Firewire/IEEE 1394 port in order to operate. ProTools is also able to operate with more than one device attached and that allows virtually an infinite number of audio channels to be processed. The mixer window consists of a left portion that presents a list of audio tracks at the top, their groups at the bottom, and the right portion delimited with a vertical split view. The right portion of the window shows visible tracks as columns for each audio track and each column corresponds to a mix table column. Each audio track consists of 5 inserts, 10 sends, input and output settings, automatic type, main volume slider, record, solo, and mute buttons, peak indicator, balance sliders and several other indicators and none of these widgets is accessible.

The mixer window is an essential portion of ProTools because it controls how the mixed data will finally appear. Each of 5 inserts presents a hardware or software DSP plug-in that can be used to filter input data such as 1 band equalizer or reverb. Inserts automatically return real-time filtered data to the same input channel. Sends are similar to inserts but do not return filtered data back and rather send it to a hardware port or a channel. Peak indicator that normally indicates too sensitive input with a red color, and record, solo, and mute buttons are also important as they change color as the only indication of being active. It is not necessary to provide access to the main volume slider and balance sliders as they copy hardware settings on the attached mix table.

The injection method is used to have bundle containing our accessorizing support loaded into ProTools' context and that bundle scans ProTools' window list and if the mixer window is not found, it installs an accessibility notification (observer) to get notified when a new window appears. Due to its special signature, mixer window can be easily recognized. Once the mixer window is available, bundle code creates `axView` hierarchy for the entire mixer window and redefines the raw keyboard down event to properly honor keyboard focus navigation. Apple advises against using the raw keyboard down event and recommends to use the unicode input text event instead so the text input methods have a chance to see the keystrokes. Hence the text input event is not sent to that window for a tab key as ProTools override it, the only way is to override the raw keyboard down event. Our event handler just scans the event for a tab key and then passes the event to the next event handler in the chain. Numerous `axViews` allow user to specify inserts and sends because there is currently no other way of changing the insert or send without a mouse. Other `axViews` provide functionality to read the name of the DSP plug-in currently inserted in given insert or send slot, or read the state of record, solo and mute buttons. Another kind of `axView` is installed on the peak indicator to notify the user when the data is at peak and might result in a data loss.

6 Conclusions

This paper provides an overview of access enabling of applications currently in development and also provides an approach to accessorize an already existing applications. The presented approach can be used to access enable Carbon applications even if they are no longer in development or its vendor has no interest in making their custom user interface elements accessible. Nonaccessible custom widgets of Cocoa applications could be accessorized with code injection too, but the mechanism would be different. The idea uses Mach kernel APIs to perform code injection that loads an accessorizing code into the application being accessorized. The code that performs access enabling has to recognize application windows and create the missing accessibility objects to enable functionality of assistive technologies such as VoiceOver.

The code injection is performed by creating the lowest level possible (Mach) thread in application being accessorized as the API that is used to create such threads allows application context to be specified. That Mach thread creates standard pthread and terminates. Pthread uses thread safe APIs to get injected code run at the system task level at which "non-thread safe" APIs can be safely called. System task level code loads accessorizing bundle that implements the missing functionality into the application either instantly or by installing accessibility observer to get notified when a new application window is available.

The accessibility is provided via `axViews` that based on `HIViews` that conform to the accessibility protocol and provide useful information to the assistive application. `axViews` are placed above the user interface elements they provide access to and can optionally monitor that elements' state to update their description and functionality. Most of `axViews` provide `kAXPressAction` action that synthesizes a mouse click on itself and as `axViews` do not handle mouse click, the click passes through and reaches the original nonaccessible user interface element.

Our approach was tested on ProTools application, that is used for professional audio editing and is mostly not accessible. The main focus was on mixer window that provides access to the mix table and the entire window was covered by a number of different `axViews` that monitor the state of the user elements and allow visually impaired used to control this window. This includes insert and send manipulation, input/output settings so the user is able to software interconnect filters implemented by DSP plug-ins with input and output.

7 Future Work

Further research can focus on similar approach that could be used with Cocoa applications and also creating an application that would serve as an accessorizing toolkit and could be used by people.

References

1. Baron, R. V., Black, D., Bolosky, W., Chew, J., Draves, R. P., Golub, D. B., Rashid, R. F., Tevanian, A., Young, M. W: Mach Kernel Interface Manual. 8/1990, Unpublished manuscript from the School of Computer Science, Carnegie Mellon University.
2. Boykin, J., Kirschen, D., Langerman, A., LoVerso, S.: Programming under Mach. Addison-Wesley Publishing Company, Inc. (1993), ISBN 0-201-52739-1.
3. Bunda, J., Potter, M. T., Shadowen, R.: PowerPC Microprocessor Developer's Guide. SAMS Publishing. (1995), ISBN 0-672-30543-7.
4. Apple Human Interface Guidelines. Apple Computer, Inc. (2005).
5. Accessibility Overview. Apple Computer, Inc. (2005).
6. Accessibility Programming Guidelines for Carbon. Apple Computer, Inc. (2005).
7. Accessibility Programming Guidelines for Cocoa. Apple Computer, Inc. (2005).
8. Engber, M.: Making Cocoa Applications Accessible. Apple Computer, Inc. (2005), WWDC 2005, Session #125.
9. Fullerton, G.: Making Carbon Applications Accessible. Apple Computer, Inc. (2005), WWDC 2005, Session #108.
10. Knaster, S.: Hacking Mac OS X Tiger. Wiley Pub., Indianapolis, IN, USA. 7/2005. ISBN 0-7645-8345-X.
11. Rentzsch, J. W.: Dynamically overriding Mac OS X.

Usability Evaluation of the MOST Mobile Assistant (SlatTalker)

Zoltan Juhasz[1], Andras Arato[2], Gabor Bognar[1], Laszlo Buday[2], Gergely Eberhardt[3], Norbert Markus[2], Emil Mogor[1], Zoltan Nagy[3], and Terez Vaspori[3]

[1] Department of Information Systems, Pannon University, Veszprem, Hungary
{juhasz, bognar, mogor}@irt.vein.hu
http://pds.irt.vein.hu/most
[2] KFKI RMKI Laboratory of Speech Technology for Rehabilitation, Budapest, Hungary
{arato, buday, markus, vaspori}@sunserv.kfki.hu
[3] Speech Technology Kft., Budapest, Hungary
{egergo, znagy}@speecht.com

Abstract. The goal of the MOST project is to develop a novel, inexpensive, easy-to-use digital talking device for blind and visually impaired users based on off-the-shelf handheld computers (Personal Digital Assistant). The device provides a novel user interface based on a simple menu system and Braille text input, and a range of application programs to support everyday tasks, including clock, notepad, phone and short messaging, email. This paper reports on the usability evaluation of the device, its strategy and implementation, and shows that our approach results in an easy to learn and use system with input speed comparable to sighted users.

1 Introduction

Miniaturisation and the consequent appearance of a range of portable digital devices (laptops, notebooks, handheld computers, mobile phones, MP3 players, sound recorders) help sighted users in performing their daily task with ever increasing efficiency. Unfortunately, this trend is not a happy one for blind and partially sighted users. For them, these devices – due to their increasingly graphical user interfaces and small control buttons – become ever more difficult to use, creating a "left out" feeling in the disappointed user. Those, however, who can overcome the hurdles and manage to use these devices, will find themselves equipped with a number of units they need to carry around.

Handheld computers (also known as Personal Digital Assistants or PDAs) would seem to be an ideal platform for creating one unified portable device that performs all the tasks of these various digital gadgets. Unfortunately, PDAs are especially not well suited to blind users for as most of the control and data input takes place via the touch sensitive screen and requires the use of a stylus.

Despite these difficulties, the MOST project is investigating how PDAs could be made usable for blind users. The aim of our effort is to develop a novel, inexpensive, easy-to-use digital talking device based on off-the-shelf PDAs that, as a result, could become the one and only device users need to carry around and yet perform a wide

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 1055 – 1062, 2006.

range of useful functions (word processing, media player, email and Internet access, mobile phone and short messaging, etc.) expected in our digital world.

The basic capabilities and functionalities of the prototype system we developed were described in a previous ICCHP article [1]. In Section 2 of this paper, we give a short overview of the current version of this prototype concentrating on the modes of operation and available functionality. Section 3 looks at the question of usability, and describes the approach we have taken to carry out usability evaluation. The primary concern at this stage was to achieve the possibility of off-site evaluation and to measure Braille input performance over a long period of time. Section 4 presents the results of our Braille input performance evaluation. The paper ends with our conclusions and directions of future work.

2 The MOST Mobile Assistant

The MOST Mobile Assistant is a special software-based solution to make off-the-shelf PDAs usable for blind users. To minimize cost, specialized hardware solutions (similar to e.g. [2]) were deliberately avoided. The navigational buttons are used to move around in the Assistant's menu structure and control programs. The touch screen contains nine virtual buttons that enable users to perform text and data input in Braille (six Braille dots, confirm button, space and backspace buttons).

Since the virtual Braille keyboard must be available at all times, a screen reader approach [3] was not suitable to make the device usable. Also, this would have locked our system to specific platforms, such as Windows or Palm. Instead, we decided to create a special software framework that provides a uniform user interface to users regardless what software they are using or who developed it.

The heart of the MOST user interface is a specific menu system that provides consistent navigation with any application. The user moves through the menu with the navigation buttons of the PDA. The *up* and *down* keys select from submenu items or commands, the *right* key executes the selected command or application, while the *left* key always returns to the previous level of the menu hierarchy.

2.1 Functionality

The approach we took required that those programs people would normally find on PDAs had to be developed from scratch. Our implementation is based on Java that makes our system platform-independent. In the prototype version we developed a number of applications. In the current state of the project, we are improving these to achieve commercial grade quality, as well as adding new programs to provide more functionality. The following programs and corresponding functionality are available at the time of writing:

- Clock: Allows users to retrieve and set current time, date and time zone. Periodic time read out at user set intervals is also possible.
- Alarm clock: Allows users to set alarm/wake-up time and corresponding sound.
- Address book: Stores user contact information such as, name, address, phone numbers, and email address. The address book can be accessed from other programs to easily retrieve information where it is needed.

- Notepad: A simple text editor that allows users to enter plain text notes or read text. Text can be read in different units (word, sentence, paragraph, continuous) as suited.
- Calculator: A simple calculator performing basic arithmetic operations. (A full blown scientific calculator has also been developed but not yet part of the system.)
- Phone: On PDA phone devices, the framework enables users to access the underlying mobile phone system to initiate and receive phone calls, receive and send short messages.
- MP3 Player: A CD quality software MP3 player that allows playback of music and talking book files stored on the PDA.
- Email Client: A special email client to receive, read and write email messages. Attachments are not supported at the moment.
- Settings: The users can also access important system settings and information. The current status of the device, battery charge level, Braille keyboard and Text-to-Speech settings can be queried and set where applicable.

2.2 Braille Input

Each of the above programs, with the exception of the MP3 player, relies on Braille input. Since the PDA touch screen senses only one pressure point at a time, it is not possible to create a software Braille keyboard in which users press several keys simultaneously similar to the traditional Braille keyboard. On the PDA, Braille input must be performed point by point, followed by a confirmation that marks the end of the Braille cell entry. For instance, the character ‘g’ is entered as the following sequence: pressing Braille dots 1, 2, 4, 5, then the Accept key.

3 Usability Testing

The primary question of the project is whether we can make PDAs usable for blind users using our approach. The functionality of the application programs has been verified previously. The number one concern is thus the effectiveness of our user interface, that is, whether the menu system is intuitive and consequent enough and the speed of entry of our Braille input method is comparable to that of sighted users.

3.1 Overview of Usability Testing Strategy

These questions can only be answered with extensive usability testing performed by blind users over a period of time. Usability tests are normally carried out in special laboratories where user activity can be closely monitored. There are several reasons why this approach could not be used in our project:

- There is no graphical user interface; traditional video recording approaches do not work.
- Testing should last for months; the number of sessions required would make testing very slow and difficult to manage.

– On-site testing would put users under unnecessary stress impairing their performance.
– Travelling to the test laboratory would impose extra burden to the users, especially for those living in different towns.

To simplify and automate usability testing and relieve users from the stress and travelling burden, a special usability logging module has been developed as an integral part of the framework. The module relies on the `log4j` logging framework [4] and allows us to record each and every event happening in the Mobile Assistant. With careful filtering it can be used to log e.g. menu navigation events, Braille dot presses, character or word entry events. This strategy lets users work at home, at any suitable time and at their own pace without stress. It also enables us to collect more reliable performance data from each user more, as logging information can be collected for several weeks or months.

3.2 Implementation

To comply with privacy and data protection directives, the logging framework is not recording every activity of the users. It is only operational in a special test program that is responsible for conducting usability tests. This test program was designed to measure the Braille input speed of the users over a long period of time. The test includes Braille character, word and sentence typing exercises during which all keystrokes and their timestamp are stored using a Java logging framework. These data provide the basis for quantitative usability analysis giving insight into how efficiently the system can be used, with what error rate and how fast the users learn to use it.

The log files generated by the usability test program are collected regularly from the users and after pre-processing stored in a database. The collected data allows us to analyse Braille input performance, such as learning rate, maximum input speed, error rate, and correlate this to computer experience, age or amount of time spent with using the device. Special care is taken to evaluate the performance of contracted and non-contracted Braille, as well as non-English character input in comparison to sighted users.

3.3 Users and Profiles

At this initial stage, five test subjects have been selected. As listed in Table 1, they are of different age groups and have a varying degree of computer experience. Some never used computers; some have experience with DOS systems, while the others are average Windows PC users.

In the future, the number of test subjects will be increased to collect more information and to find out how younger adults and children perform on the Mobile Assistant.

Several sighted users acted as control group test subjects (age 26-40) in order to establish baseline performance for normal stylus-based PDA entry speed. These users performed the sighted version of the usability test program.

Table 1. The profiles of the test subjects in terms of age, sex, sight impairment and computer experience

User No.	Gender	Age	Degree and age of sight loss	Use of Braille	Computing experience
1	female	63	blind, age 3	regular, contracted	none
2	male	54	blind, since birth	regular	average user
3	male	42	blind, since birth	regular, contracted	expert user
4	female	45	blind, age 11	regular	average user
5	male	42	blind, age 17	seldom	expert user

4 Results

In this paper we discuss results related only to the Braille input speed measurement, as we consider this the most important factor in the usability of the system. From the Braille input point of view, there are two key usability-related characteristics that determine the viability of our approach. The first is whether blind users are able to use the touch screen-based Braille keyboard. The second is the achievable typing speed.

We have found that although it is possible to use the touch screen without a positional guide, this way of operation results in very slow input and high error rate. To orient users, a plastic guide with holes at the dot locations have been designed that can be fitted onto the PDA in a way that it does not touch the screen. The user then only needs to find the particular hole for the given virtual key and the hole will guide his or her finger to the correct touch position

From the test logs we can determine the amount of time spent with the usability tests as well as with other activities. This provides important information about how much experience the user gathered during the evaluation period. Table 2 summarises the tests quantitatively. During the entire test period, our blind users typed several hundreds of character a day, in total varying between 12 000 and 27 600. The log files also indicated that on average, users 1-5 used the Assistant for 121, 50, 71, 89, 156 minutes each day, respectively.

Table 2. The average time spent with daily testing, and the number of tests performed

User No.	Active testing days	Daily average			
		Testing time [min]	Braille character tests	Word tests	Sentence tests
1	27	38	44	26	10
2	15	21	31	21	4
3	47	18	54	25	5
4	69	25	50	25	7
5	34	41	23	36	19

4.1 Typing Speed

The data in the log files allowed us to explore in detail how users can use our system. We could extract information such as the time to enter Braille characters, complete words and sentences; the variation of these time values with time that show speed improvement due to learning; the type and rate of errors during text input; correlation between age and entry speed, and correlation between entry speed and the number of Braille dots required for the text.

For space considerations, we only show one set of results in Fig. 1 and 2, which summarize the character entry time distributions for the blind versus sighted users and show the variation in error rate with time.

Based on the app. 150 msec response time of PDA screens and the average user reaction time, we presumed that 2 virtual buttons can be pressed in a second on the touch screen. Taking an average of 3 Braille dots per character plus the Accept key press, this results in 2-second average character entry times, or 30 char/min typing speed. Fig. 1 clearly shows that 3-point characters are entered faster than anticipated, in app. 1.5 seconds. In fact, nearly all characters are typed in less than 2 sec average.

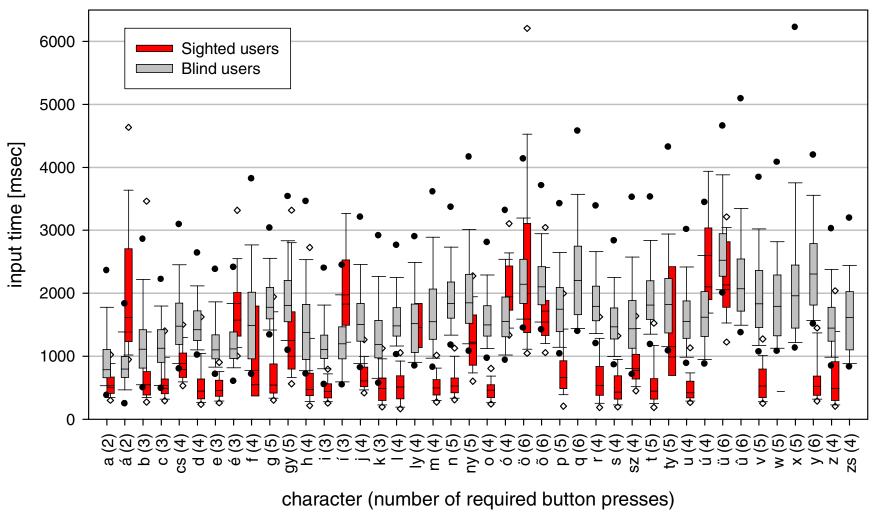


Fig. 1. Box plots of blind and sighted user input times. Boxes represent boundaries for 25-75% of data points with median, whiskers for 10-90%, and outliers for 5-95%.

When compared to the entry times of young (age 25-35) sighted users, we can see that the English alphabet characters are entered 2-3 times slower. (Note, that this speed difference is expected to disappear for older sighted users.) Characters with accents (e.g. á, é, ű) or multiple consonants (e.g. ny, sz, ty), however, are typically entered faster by the blind users on our Braille keyboard. Word and sentence test also show that blind users achieved half of the entry speed of young sighted users. We expect to see further speed gain when using contracted Braille input.

Figure 2 illustrates the variation of the relative error (number of incorrect characters / total typed characters) in Braille text entry. The results demonstrate how fast users can learn to use the Mobile Assistant. Users, who never before used a PDA device and our Mobile Assistant, could naturally and quickly adapt their Braille skills to the Braille touch screen keyboard. The error curves show that within two weeks each user converged to app. 1% error from the 5-10% initial value.

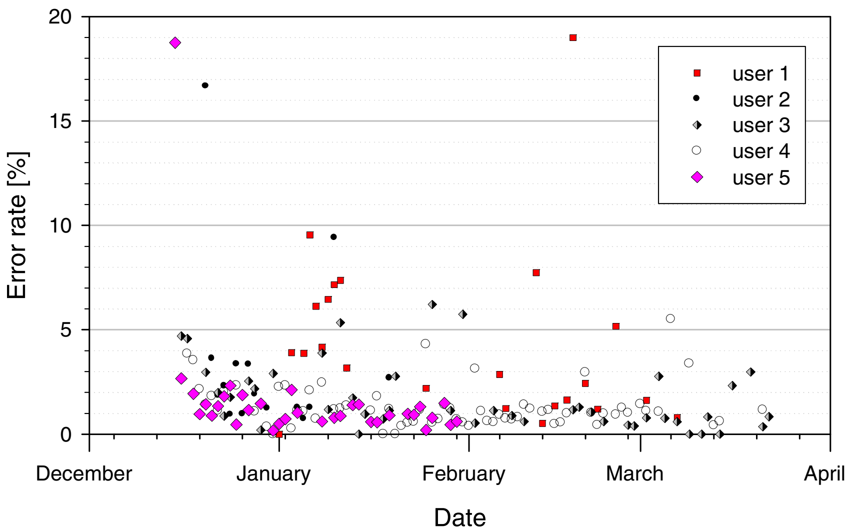


Fig. 2. The change of error as users spent more time using the Mobile Assistant. Each user achieved an app. 1% error rate within two weeks of testing.

5 Conclusions

The aim of the MOST project is to create an affordable and efficient portable computing device based on off-the-shelf PDAs. The purpose of this paper was to give an overview of the MOST Mobile Assistant, its functionality and user interface, and describe in detail the approach we took for the usability evaluation of the device.

We have developed a special logging module that allows users to perform usability tests without stress at any suitable time in their home. The framework provides detailed information about the performed usability tests for a long period of time. The evaluation of the results showed that text entry speed comparable to sighted PDA users can be achieved and users learn to use the system very quickly. This is very important as we intend to create an Assistant that helps users who do not have computing knowledge.

Further input speed improvement is expected by the use of contracted Braille input or external Bluetooth keyboard that provides sighted typing speed for blind users with standard keyboard typing skills. These tests are being planned at the time of writing and should produce results in the near future. Further usability tests are also planned that will involve more users including children. Besides pure text entry measurements, task-oriented performance studies are planned to measure e.g. the effectiveness of sending short messages from the Assistant in comparison to sighted PDA and mobile phone users.

A set of new application programs is also under development. An agenda program, simple web browser, colour recogniser, educational, adventure and quiz games are in the pipeline. Work is underway to create a client application to access the Hungarian E-book Library [5] and other public electronic libraries for browsing, searching and downloading books (in either text or mp3 format) in a uniform way.

In the long term, the Mobile Assistant will be ported to desktop computers to provide blind users with a simpler and more effective environment than today's graphical user interface oriented systems.

Acknowledgements

The authors would like to acknowledge the support of the National Office for Research and Technology for their support under Grant No. GVOP-3.1.1-2004-05-0266/3.0. We would also like to thank for the effort of our undergraduate students in performing smaller programming and evaluation tasks.

References

1. Arato, A., Juhasz, Z., Blenkhorn, P., Evans, G., Evreinov, G.: Java-powered Braille slate talker. In: Proc. ICCHP 2004: 9th International Conference on Computers Helping People with Special Needs, July 7-9, 2004, Paris, France, Lecture Notes in Computer Science, Vol. 3118, pp 506-513, (2004)
2. Freedom Scientific Inc.: PAC Mate BNS, (2004) http://www.freedomscientific.com/fs_products/PACmate_bns.asp
3. Technologies HumanWare Canada Inc.: Maestro PDA, http://humanware.ca/web/en/ p_DA_Maestro.asp, (2005)
4. log4j project: http://logging.apache.org/log4j/docs/
5. Hungarian Electronic Library: http://vmek.oszk.hu/

Non-visual Access to GUIs: Leveraging Abstract User Interfaces

Kris Van Hees and Jan Engelen

Katholieke Universiteit Leuven
Department of Electrical Engineering
ESAT - SCD - DocArch
Kasteelpark Arenberg 10
B-3001 Heverlee, Belgium
kris@alchar.org, jan@docarch.be

Abstract. Various approaches to providing blind users with access to graphical user interfaces have been researched extensively in the past 15 years, and yet accessibility is still facing many obstacles. Graphical environments such as X Windows offer a high degree of freedom to both the developer and the user, complicating the accessibility problem even more. Existing technology is largely based on either a combination of graphical toolkit hooks, queries to the application and scripting, or model-driven user interface development. Both approaches have limitations that the proposed research addresses. This paper builds upon past and current research into accessibility, and promotes the use of abstract user interfaces to providing non-visual access to GUIs.

1 Introduction

Ever since graphical user interfaces (GUIs) emerged, the community of blind users has been concerned about its effects on computer accessibility [1]. Until then, screen readers were able to truly read the contents of the screen and render it in tactile and/or audio format. GUIs introduced an inherently visual interaction model with a wider variety of interface objects. MS Windows rapidly became the *de facto* graphical environment of choice at home and in the workplace because of its largely consistent presentation, and the availability of commercial screen readers created somewhat of a comfort zone for blind users who were otherwise excluded from accessing GUIs.

As X Windows-based systems grow in popularity for both home and work environments, an additional level of complexity has emerged. Not only pose GUIs an obstacle by being inherently visual, but it is also possible to combine elements from a variety of graphical toolkits such as Athena, GTK, Qt, … into a single graphical environment. This important feature promotes flexibility and interoperability, but it largely complicates the work needed to provide accessibility. Screen readers either must support all commonly used toolkits, or they must be designed to not depend on any implementation specific details.

Past and current research indicates that abstracting the user interface offers a high degree of flexibility in rendering for a multitude of output modalities. Blind users generally prefer auditory and/or tactile representations of the GUI, as an alternative to the visual rendering provided by GUIs. Current approaches use a combination of toolkit

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 1063–1070, 2006.

extensions, scripting, and complex heuristics to obtain sufficient information to make alternative renderings possible [2]. Alternative approaches aim to solve the accessibility problem by addressing the different output modalities in the design and development of applications, composing model-driven user interface implementations at development time.

The need for alternative user interface representations across different output modalities makes the accessibility problem a prime candidate for using abstract user interface (AUI) descriptions. The remainder of this paper first presents related work on GUI accessibility. The third section describes the use of AUIs at the core of an accessibility framework, providing non-visual rendering in parallel with visual representations, followed by a comparison of this approach against AUI-based model-driven user interface construction. Section five concludes this paper with a description of future work.

2 Related Work

In the context of accessibility of GUIs for blind users, Mynatt and Weber provided two early approaches [3]. The Mercator project at the Georgia institute of Technology replaced the GUI with a hierarchical auditory interface, whereas GUIB provided a tactile representation of the screen contents. Contrasting both projects also established four core design issues that are common to non-visual access to GUIs, further refined in [4]. This early paper also provides a more general description of common approaches towards GUI accessibility.

The ”Fruit” system described by Kawai, Aida, and Saito [5] addresses the issue of user interface accessibility by means of an abstract widget toolkit. Application software is still written as if a graphical widget toolkit is being used, while the actual presentation of the user interface is handled by device-specific components. The ”Fruit” system does not support synchronised presentation in multiple modalities, nor does it provide any accessibility at the level of the windowing environment.

Savidis and Stephanidis explored alternative interaction metaphors for non-visual user interfaces [6]. This work was expanded upon in the development of a user interface development toolkit [7]. The HAWK toolkit provides interaction objects and techniques that have been designed specifically for non-visual access. The toolkit is used in the AVANTI project [8], introducing a Unified User Interface concept using runtime user interface adaptation based on user and usage context.

The similarities between application user interfaces and World Wide Web forms provide an important foundation for using AUIs. Barnicle researched specific obstacles that blind users encounter when using GUIs [12]. His results were confirmed in later research[13,14]. The UsiXML [9] project at the Belgian Laboratory of Computer-Human Interaction (BCHI) at the Université Catholique de Louvain builds upon these concepts. The ability to abstract the user interface lies at the core of the research presented in this paper.

Research into alternative user interfaces has been extensive. Two notable projects are the ”virtual sound wall” at the Oldenburg R&D-Institute for Informatics Tools and Systems [10], and the performance analysis of multi-modal interfaces by Vitense, Jacko, and Emery [11]. Both solutions require expensive devices that are well outside the budget of an average user. The research presented in this paper therefore limits the context of output modalities to refreshable Braille displays, speech synthesisers, and/or non-spatial sound.

3 Leveraging AUIs Towards Accessibility

The non-visual presentation of GUIs poses several HCI design issues that need to be addressed as part of any acceptable accessibility solution. Five fundamental problems were identified by Mynatt and Weber [3], and Gunzenhäuser and Weber [4]:

- Coherence between visual and non-visual interfaces
- Exploration in a non-visual interface
- Conveying graphical information in a non-visual interface
- Interaction in a non-visual interface
- Ease of learning

Abstract user interface rendering at runtime ensures coherence, and allows both sighted and blind users to operate using the same mental model of the interaction semantics [15,16]. In addition, AUIs can address the three remaining fundamental HCI problems by translating what is perceived as visual metaphors in a representation independent manner. Some of this task is to be delegated to the screen reader implementation. E.g. while the AUI rendering is handled per application by shared components, the screen reader is typically implemented as an application that serves the entire windowing environment. Figure 1 explains how non-visual access to GUIs can be implemented by leveraging AUIs.

Whereas current approaches work with applications that are implemented against a specific graphical toolkit, depending on explicit support for accessibility in the toolkit,

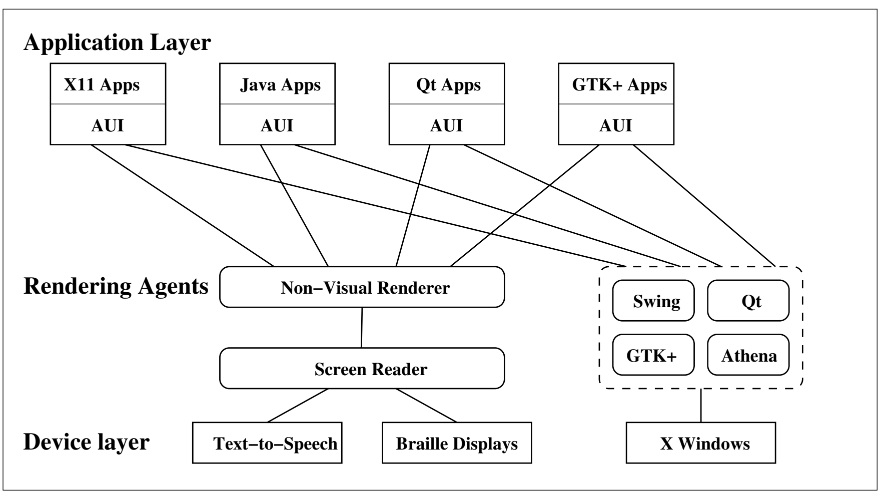


Fig. 1. Leveraging abstract user interfaces for accessibility

AUI-based non-visual access builds on a paradigm where the UI of the application is described in an abstract form. The presentation of the UI is delegated to specific rendering agents, both visual and non-visual [15].

The advantages of this approach are not only within the context of providing non-visual access to GUIs but also within the context of user-controlled “look & feel”. Because applications are no longer specifically built for a given graphical toolkit and rendering is delegated to specific agents, it is possible to render applications against any supported toolkit. This flexibility can be a powerful feature for many users.

Leveraging AUIs allows accessible user interfaces to be implemented side-to-side with their graphical counterparts. This resolves a long standing problem with screen readers needing to tap into the application flow in order to retrieve the information needed to drive alternative presentations. This puts current screen reading technologies at a definite disadvantage.

The XML document shown in figure 2 provides an example of a fairly simple user interface. The graphical rendering in Swing is shown in figure 3. The example application is a card game, appealing to both a sighted and blind audience.

While the example presented here is very basic, it does show important features of the presented approach. The menu bar with all its components is defined in a truly abstract way, whereas the main user interface uses an abstract form augmented with additional (graphical) information, such as image sizes and background image specifications. It carries all information to facilitate both visual and non-visual presentations. While the example is a hand-crafted AUI description, more complicated applications would benefit from generated user interface descriptions (e.g. using UsiXML [9]).

The reference implementation currently being developed parses the XML document describing the user interface into an abstract object model. This model drives all

```
<?xml version="1.0"?>
<gui>
  <window id="GameSys">
    <menuBar>
      <menu id="Game">
        <menuItem id="New" label="New game"/>
        <menuSeparator/>
        <menuItem id="Quit"/>
      </menu>
      ...
    </menuBar>
    <form id="Table" width="500" height="300"
          bgImage="felt.jpg">
      <form id="Card-1" width="71" height="96"
            bgImage="card-1.gif"/>
      ...
    </form>
    <statusBar label="GameSys v1.0.0"/>
  </window>
</gui>
```

Fig. 2. XML document providing an abstract user interface description

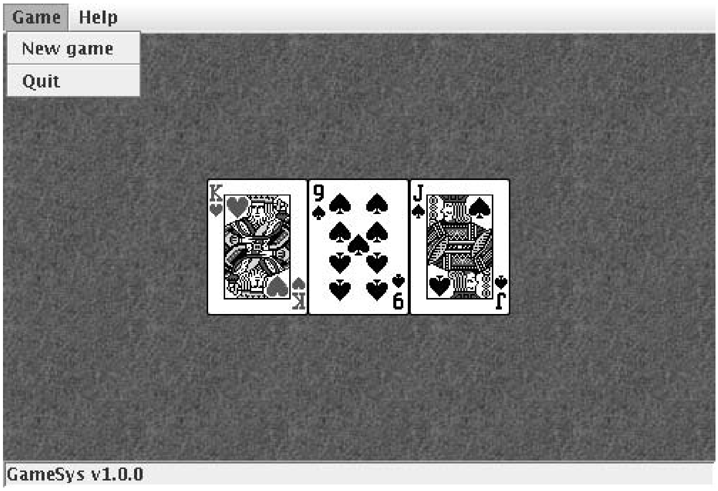


Fig. 3. Rendering an AUI description using Swing

rendering engines, providing them with information about what needs to be communicated to the user. The rendering engines decide how the information is presented. Input from the user is channelled from the input device to the focus manager in the abstract object model, interacting both with the application and with the rendering engines. The current architecture for the prototype is shown in figure 4. Note that whereas the XML document describes only how the user interface is presented to the user, the abstract object model also carries information on what data is to be presented as is suggested by earlier research [17].

Although Swing is used as the underlying technology for this prototype, all of the UI logic is handled by the AUI focus manager. This ensures that both the visual and the non-visual presentations provide users with identical interaction models. In addition, the focus manager can make decisions based on graphical information (element sizes, etc…) without exposing the actual decision logic to the rendering agents. This is required in order to resolve the problem of conveying graphical information in a non-visual interface.

4 Comparison with Model-Based UI Construction

Model-based UI construction generally covers two approaches:

– Development-time UI construction
  This includes all ways of generating multiple final user interfaces (generally for a variety of modalities). Often, rather than simply creating multiple UI front-ends, multiple versions of the application are generated due to using a more invasive interface between the application and the user interface. A prime example is the UsiXML project (and its derived projects) [9].

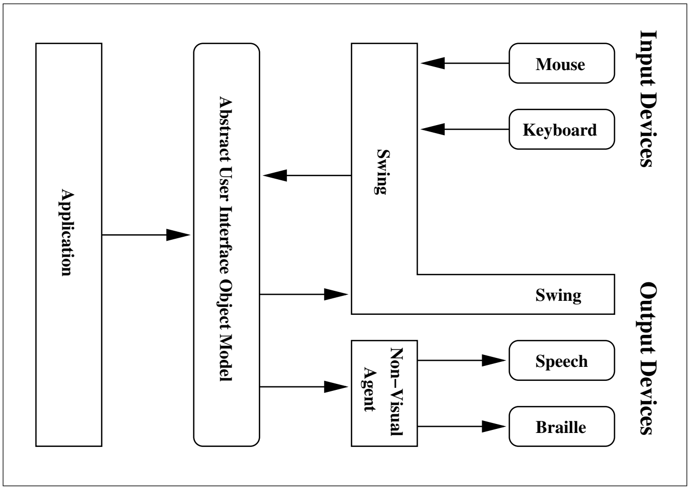


Fig. 4. Reference implementation architecture

– Runtime UI selection
  This includes all mechanisms that allow some form of model-driven UI customisation. Where development-time UI construction generates a set of UIs, runtime UI selection supports alternatives for UI components. A model-driven UI engine determines which alternative (if any) for a given component is to be used at any given time. A good example of this approach is the AVANTI project [8].

The research presented in this and earlier papers [15,16] can and will build on development-time UI construction techniques, but it present a very different approach in terms of accessibility. Providing non-visual access to UIs by means of modality specific implementations constructed during the development of the application makes it impossible to share a single instance between a sighted and a blind user. As such, collaboration by observing the same runtime information is not possible. Runtime rendering of the AUI allows for this by supporting multiple simultaneous renderings.

Runtime UI selection is somewhat similar to the approach presented here, in the sense that the user interface presentation is decided upon at runtime rather than at development time. This powerful technique supports current concepts of universal access, and is a major advancement towards accessibility. It does not generally provide for simultaneous rendering in different modalities. It does involve a tight coupling between the application and the user interface presentation, making future extension of supported modalities more complicated. Runtime rendering of AUIs does not have that limitation.

5 Conclusion

This paper describes a novel approach to providing a long-term solution for non-visual access to GUIs by leveraging abstract user interfaces. The theory behind the presented work is built upon extensive research into HCI accessibility issues, AUIs, past and current approaches to solving this complex problem, and user feedback on existing implementations.

A prototype is being developed for extensive field testing as part of this research. Only through user feedback can success be measured appropriately. The prototype will be enhanced to provide better support for a representative subset of commonly used widgets so that multiple example applications can be built.

There is still a long way ahead, and ultimately, integration with existing (and in-development) AUI frameworks will be important. Not only because modifications may be required in order to support runtime rendering and non-visual access, but also because the proposed approach builds upon the adoption of the AUI application development paradigm. because

Acknowledgements

The research presented in this paper is part of the author's doctoral work at the Katholieke Universiteit Leuven, Belgium, under supervision by Jan Engelen (ESAT-SCD-Research Group on Document Architectures).

References

1. L. H. Boyd, W. L. Boyd, and G. C. Vanderheiden. The graphical user interface: Crisis, danger and opportunity. In *Journal of Visual Impairment and Blindness*, pages 496–502, 1990.
2. Gerhard Weber and Rolf Mager. Non-visual user interfaces for X Windows. In *Interdisciplinary aspects on computers helping people with special needs*, pages 459–468. 5th International Conference, ICCHP 96, 1996.
3. Elizabeth D. Mynatt and Gerhard Weber. Nonvisual presentation of graphical user interfaces: Contrasting two approaches. In *Human Factors in Computing Systems*, pages 166–172. CHI 94 – Celebrating Interdependence, 1994.
4. Rul Gunzenhäuser and Gerhard Weber. Graphical user interfaces for blind people. In Brunnstein K. and E. Raubold, editors, *13th World Computer Congress 94, Volume 2*, pages 450–457. Elsevier Science B.V., 1994.
5. Shiro Kawai, Hitoshi Aida, and Tadao Saito. Designing interface toolkit with dynamic selectable modality. In *Proceedings of the second annual ACM conference on Assistive Technologies*, pages 72–79. ACM Press, 1996.
6. Anthony Savidis and Constantine Stephanidis. Building non-visual interaction through the development of the Rooms metaphor. In companion of the *CHI'95 conference in Human Factors in Computing Systems*, pages 244–245, 1995.
7. Anthony Savidis, Athena Stergiou, and Constantine Stephanidis. Generic Containers for Metaphor Fusion in Non-Visual Interaction: the HAWK Interface Toolkit. In *Proceedings of the Interfaces '97 Conference*, pages 194–196, 1997.
8. Constantine Stephanidis and Anthony Savidis. Universal Access in the Information Society: Methods, Tools, and Interaction Technologies. In Reinhard Oppermann, editor, *Universal Access in the Information Society*, 1(1):40-55, 2001.
9. Q. Limbourg, J. Vanderdonckt, B. Michotte, L. Bouillon, and M. Florins. Usixml: A user interface description language supporting multiple levels of independence. In M. Lauff, editor, *Proceedings of Workshop on Device Independent Web Engineering DIWE'04 (Munich, 26-27 July 2004)*, 2004.

10. Hilko Donker, Palle Klante, and Peter Gorny. The design of auditory user interfaces for blind users. In *Proceedings of the second Nordic conference on Human- Computer Interaction*, pages 149–156. ACM Press, 2002.
11. Holly S. Vitense, Julie A. Jacko, and V. Kathlene Emery. Multimodal feedback: establishing a performance baseline for improved access by individuals with visual impairments. In *Proceedings of the fifth international ACM conference on Assistive Technologies*, pages 49–56. ACM Press, 2002.
12. Kitch Barnicle. Usability testing with screen reading technology in a windows environment. In *Proceedings on the 200 conference on Universal Usability*, pages 102–109. ACM Press, 2000.
13. E. Pontelli, D. Gillan, W. Xiong, E. Saad, G. Gupta, and A. I. Karshmer. Navigation of html tables, frames, and xml fragments. In *Proceedings of the fifth international ACM conference on Assistive Technologies*, pages 25–32. ACM Press, 2002.
14. Mary Frances Theofanos and Janice Redish. Bridging the gap: between accessibility and usability. *Interactions*, 10(6):36–51, 2003.
15. Kris Van Hees and Jan Engelen. Abstract UIs as a long-term solution for non-visual access to GUIs. In *Proceedings of the 11th International Conference on Human-Computer Interaction* (CD-ROM), 2005.
16. Kris Van Hees and Jan Engelen. Abstracting the Graphical User Interface for Non-Visual Access. In A. Pruski and H. Knops, editors, *Assistive Technology: From Virtuality to Reality (AAATE 2005)*. IOS Press, 2005.
17. Shari Trewin, Gottfried Zimmermann, and Gregg Vanderheiden. Abstract user interface representations: How well do they support universal access? In *CUU '03: Proceedings of the 2003 conference on Universal usability*, pages 77–84. ACM Press, 2003.

Design Guidelines for Audio–Haptic Immersive Applications for People with Visual Disabilities

Fernando Alonso[1], José L. Fuertes[1], Loïc Martínez[1], and Héctor Szabo[2]

[1] Facultad de Informática, Universidad Politécnica de Madrid, Campus de Montegancedo, 28660 - Boadilla del Monte, Madrid, Spain
{falonso, jfuertes, loic}@fi.upm.es
[2] Facultad de Ingeniería, Universidad ORT Uruguay, Cuareim 1451, 11100 Montevideo, Uruguay
szabo@ort.edu.uy

Abstract. We describe the goals, issues and constraints found creating audio-haptic applications oriented to people with visual disabilities. Using audio-haptic games as a measuring tool, we have found unexpected features in the user's audio perception. Our goal is to define a model of perception and usability guidelines for developers creating immersive accessible applications. This paper presents initial findings, related to user precision over 3D audio and haptic effects. The game and environment use a multichannel speaker array taking advantage of Microsoft DirectX multichannel audio support (5.1) as the audio processing abstraction layer. The environment allows simple haptic assistance through force feedback joystick devices, also supported by Microsoft DirectX as the force feedback and user input abstraction layer.

1 Introduction

We have been developing accessible games since 2003 at ORT University, Uruguay, as testing tools for immersive audio–haptic environments, based on previous audio game concepts [1,2]. These games used 3D audio with simple haptic support. We used Microsoft DirectX as the hardware abstraction layer for game development [5]. We handled multichannel 5.1 audio by means of MS DirectX 3D acoustic support (DirectSound) [6]. In addition, we made use of basic haptic feedback with MS DirectX force feedback support (DirectInput) and a force feedback joystick. Force feedback joysticks are simple haptic devices and provide basic features: haptic exploratory procedures supported are pressure and contour following (with limitations), according to Klatzky's classification [7]. We did a proof of concept for the environment and gathered preliminary feedback from users of the 2003 game, augmented by new opinions and perceptions in 2004. During 2004, we developed an "audio-haptic engine" called "SHADE" (Simple Haptic Acoustic Development Engine) [8], which facilitates game production and simplifies the way in which more sophisticated effects can be used. During 2004 and 2005 we used SHADE to develop games [10].

During the testing and concept validation phases in 2003, we found that some users' 3D sound perception differed from the expected behavior [4,9]. To establish a set of predictably perceived sound and haptic effects, we started to create new SHADE-based games to automatically measure and record user responses and their subjective

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 1071 – 1078, 2006.

overall assessment for a selected effect set. We expect to begin testing these games with end users in 2006.

In the following sections, we will describe observed user perceptions in our present research and future lines of research. Section 2 describes the test environment and explains SHADE and DirectSound capabilities and constraints. Section 3 explains validation test findings. Section 4 gives a preliminary limited set of guidelines for developers working with multichannel audio–haptic environments. Section 5 sets out our preliminary conclusions and our expected future lines of research.

2 SHADE Aural-Haptic Display

SHADE strongly relates to 3D DirectSound and DirectInput models. The DirectSound 3D model defines, for each sound source, its volume level, direction, and positioning in the 3D virtual soundscape. The sound model assigns a virtual position to the user in a 3D space. The x-axis goes from left to right; the y-axis goes from bottom to top; the z-axis goes from near to far [6]. Sound source positioning may be absolute or relative to the avatar[1]. Sound propagation rules may establish two "sound projection cones" (Figure 1) and special effects like "occluded sound" (behind a wall or door).

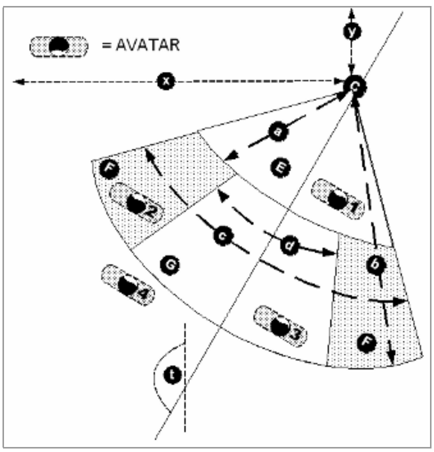


Fig. 1. DirectSound propagation model, showing inside cone *(d)*, outside cone *(c)*, and virtual zones with particular sound propagation. Distance *(a)* represents the "near" zone, distance *(b)* represents the sound fading zone and no sound is perceived over distance *(b)*. The sound source is located at the origin (C). [2]

Looking at Figure 1, when the avatar's virtual position is in zone *(1)* (marked over the avatar's right shoulder), it is meant to be near the sound source, and all speakers have the same volume level and play synchronously. Therefore, the sound source

[1] Avatar: from Sanskrit *avatâra*, god's incarnation on earth. Term used to refer to the main character in role-playing games; usually the player "tells" this character what to do to play the game.

position is not identifiable. When in zone *(2)*, the avatar is meant to be far away from the sound source and only the secondary cone sound (strongly faded) is perceived. When in zone *(3),* the avatar is meant to be in a direct path, but far away from the sound source. Therefore, the user clearly perceives the sound, but its volume is faded. When the avatar is in zone *(4),* it is meant to be so far away from the sound source that it either perceives no sound at all or the sound volume is very low and constant. In all these cases, differential time shifting is applied to each loudspeaker [6].

Although we can theoretically change elevation (vertical position) using DirectSound to shift the sound timing, it is very hard to get different users to have a stable perception of elevation [6,9]. This is because people perceive elevation cues as a natural time shift produced by the outer ear (pinna), person's height and other morphological aspects [11]. For the time being, then, we have opted for a 2D soundscape in SHADE, because of the above-mentioned instability and the increased complexities in loudspeakers setup (loudspeaker position and room height, among others) [9].

Our preliminary approach to 3D sound-haptic games intended users to sense an immersive environment and to improve on traditional stereo or monaural audio game playing experiences. User feedback indicated satisfaction with the environment's increased realism and complex capabilities offered. We expect people to use this environment at home; as a result, standard off-the-shelf equipment was to be used.

We decided not to produce a graphical interface for SHADE-based games to make sure that both sighted and blind users had the same clues. We plan to add high-contrast wire frame 3D scenes for low vision users in future releases.

Joystick haptic capabilities offer limited haptic support, including slow motion (0.2 Hz) (i.e., simulating waves in a calm sea) or fast motion (500 Hz) (i.e., simulating a gong), in which case amplitude must be short. It can also rotate or simulate a quick movement, like the release of a bow string (from fast to slow movement) or some object being hit (from slow to fast movement with an abrupt end). Joysticks can only vibrate and rotate with limited amplitude and frequency [13] and cannot offer real enclosure or full contour sensation. Other products, like PHANToM from SensAble Technologies [14] or the SPIDAR from Tokyo Institute of Technology [15], among others, offer richer support, but are experimental or less accessible to targeted end users. When we began our work in 2003, we also found that the force-feedback joystick was very similar to a very old and common haptic device for blind people: the white cane. Considerable work has been done worldwide using the force-feedback joystick as a haptic device since 2002 and information is still being gathered [4].

3 Validation Test Findings

3.1 Analyzed Variables

We evaluated perception accuracy in terms of the angular distance between the theoretical sound source position and the user signaled position, following previous immersive sound experiments conducted by the CIPIC Interface Laboratory [9,11]. Even though we are working in a 2D sound space only, this space is part of a 3D environment. Therefore, we still refer to environment sounds as 3D sounds.

Working with SHADE and DirectSound, we defined the term "*3D Sound Azimuth*" (3DSA) for the user perceived angular position of a 3D sound source. 3DSA is the avatar's angular orientation in relation to an arbitrary "front" position in the soundscape (referred to as angle 0º "Front", see Figure 2b) and the virtual 3D sound source position with respect to this.

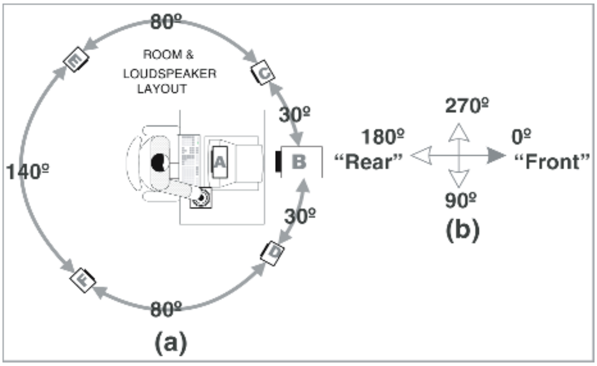


Fig. 2. (a) Loudspeaker array with user [17]. (b) Arbitrary angles assumed in environment evaluation.

We will use the term *3DSA Accuracy* (3DSAA) to indicate the accuracy of the user perception of the 3D sound position. This metric is defined as the angle between the extreme azimuth perceptions for the same sound position. Lower angular values represent better 3DSAA behavior, i.e., a 3DSAA of 30º means an accuracy of ±15º from the sound source position.

Our initial observations related to 3DSA and 3DSAA using low (30 Hz) and high (3000 Hz) audio frequencies. Our sound probes were "contextual" sounds: a water drop, a cricket, and thunder, among others. We observed user reaction in a small group of people who participated in product tests: five sighted people and three visually impaired users who were 16, 23 and 37 years old at the time they tested the products. We also evaluated haptic support. Blind people found it too strong, but still considered it to be interesting and to reinforce game play.

3.2 User Perceptions

How accurate is user 3DSA appreciation? Is it stable at different sound frequencies? Can the end user perceive that a sound source is moving virtually around the avatar's position?

These questions arose when we began to test our game and found that some of the user perceptions differed from expected behavior. These irregularities were also found to occur in other games.

Currently, we have no systematic or automatic records, because earlier games did not collect user reaction time. We expect to automate information gathering in two new games to be delivered in early 2006. However, taking in account our initial data and subject to future validation, a short summary of the main perception problems detected is presented below.

1. **Effect of sound source azimuth** (sound position effect). Primary sound azimuth cues are binaural [9], [11], [12]. For static sources, there is a significant variation in

user 3DSAA depending on sound position. Perception quality increases for lateral azimuths (more exactitude in angle appreciation) and strongly decreases behind the user (Figure 3a). We can see that users have a 3DSAA of 45º with frontal audio sources and 3DSAA quality increases to 30º (±15º) for lateral sound sources. Finally, we find that there is a sharp loss of precision for sound sources behind the user. This demonstrates the effects of Interaural Time Difference (ITD)[2], which is at its maximum level when the environment produces the sound laterally and drops to a minimum when the sound is produced in front of the listener [5]. Behind the user, outer ears and head "shadow" stop the sound being heard directly: complex sound phase shifting and room reverberation alters perception, making it even more difficult to identify the source's 3DSA [4].

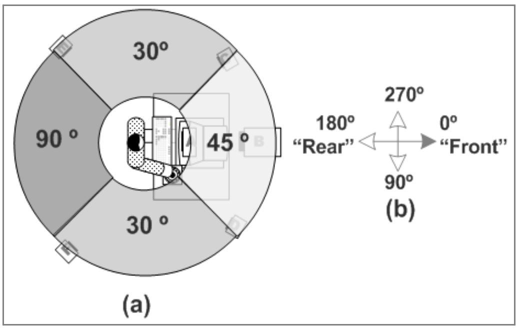


Fig. 3. (a) Mean observed 3DSAA among environment testers. (b) Arbitrary angles assumed in environment evaluation.

2. **Effect of sound source relative distance** (far–near effect). Due to the DirectX model of sound propagation and *contrary to* expected behavior, 3DSAA is better for virtually distant sources than for nearer ones. Sound sources virtually located at 3 meters from the avatar have a lower 3DSAA (greater angle) than sound sources virtually located at 5 meters or more.
3. **Effect of sound source frequency.** At lower frequencies, the sound wavelength is larger than head diameter and the room reverberation effect is worse [12]. We found that user azimuth identification was less reliable with low frequency sources. Sometimes they could vaguely identify what "regions" the sound came from ("from the front", "from the left", and so on), whereas on other occasions they could not even identify a region, especially when low frequency noises were "close" and the room reverberation had a worse effect on perception. The use of one subwoofer as the only low frequency emitter also affects results: low frequency sounds are only emitted by the subwoofer in position "B" (Figure 2a). The combined effect of low frequency and rear sound position caused a 3DSAA as bad as 120º for some users.
4. **Effect of "ambiance" sound incidence.** We used two kinds of "theatrical ambiance" sounds: non-directional thunder and automatic characters: a dog (barking)

[2] Interaural Time Difference: Delay time between the signal perceived by one ear with respect to the other. Its value is highest when the sound source is in a lateral position and the difference in the ear distance with respect to sound source is the greatest.

and guards. Thunder claps were disorienting as users could not define the 3DSA and expected some clue from the sound. The barking dog sound made one of our users jump: she believed that there was a real dog in the room. In this case too, users expected some clue from those sounds.

5. **Effect of inner-outer cone sound volume ratio.** We observed a 3DSAA quality reduction when outer cone sound volume exceeded inner sound volume by 30%.
6. **Effect of sound motion.** It is difficult for users to perceive small changes in 3DSA. Variations under 10º were undetected for all users and more than 70% of users could not detect 3DSA changes of less than 45º (see Figure 4).

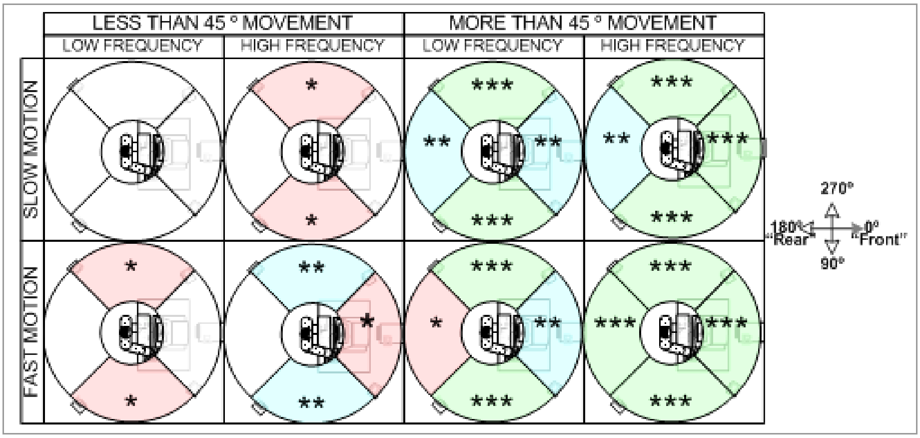


Fig. 4. Average user ability to detect 3DSA motion. (***): all test users, (**): 50% of test users, (*): 25% test users, ( ): none. User is looking to the right in this diagram.

7. **Effect of haptic reinforcement.** Haptic support reinforces 3DSA perception, but does not increase precision. Users like haptic effects as clues or theatricals reinforcement. We need to conduct experiments over a bigger population to develop guidelines for haptic features.
8. **Effect of user training.** Although we acknowledged a 3DSAA enhancement for users playing the game more than once, more studies need to be undertaken to get a reliable quantitative measure.

4 Initial Developer Guidelines

This section presents our preliminary guidelines for developers, which we expect to extend after further work.

The more challenging an effect is, the fewer the people who will be able to perceive this effect. We intend to classify user audio-haptic capabilities into different skill levels. In this way, categories, i.e., "beginner", "novice" and "expert", will be able to be defined for expected skills needed for use of simulators and games or to be able to do jobs using an auditory display. We also expect to define a usability/playability guide for developers, based on these skills levels.

1. Offer a high pitch "reference sound" in each environment. This allows users to get their bearings in the soundscape and to construct the concept of "front", "right", "left" and "rear", simplifying their virtual movement interpretation.
2. Use at least 45º in movement implementation. This allows most users to recognize soundscape rotation. In SHADE, the avatar is only allowed to perform 90º turns.
3. Use low frequency sounds (less than 100 Hz) for omni-directional background sounds. Use them only if they are periodic and well documented (i.e., for thunder, the initial audio-drama explains that it is a stormy night). Do not use low frequency sounds as directional guides.
4. Use wide movements to move objects in the soundscape. If this is impossible, use only high-pitched sounds to refer to these objects in the soundscape.
5. Allow objects to refer (announce) themselves from long distances; this will make user orientation more precise. When the user comes close to the object, there is a loss in precision. Compensate for this loss using a secondary distant referrer sound in another position, preferably at 90º from the target sound source.
6. Rear sounds are indicative of presence and are not precise. Give users enough time to rotate to raise precision before user interaction is needed.
7. Design initial clues as high-pitched (2000–5000 Hz) sounds. A few minutes into the game, users may get their bearings with lower-pitched sounds, but designers should refrain from using very low frequencies (less than 100 Hz) for actions requiring orientation and precision.
8. Reinforce low-pitched clues with haptics indicating direction, if possible. Haptic scores do not rotate around the joystick axis (their angle is fixed), which is another good reason for allowing only very limited angle positions.
9. Use small amplitude movements with haptic reinforcement: blind users will find it less exaggerated, and all users will still feel the score.

5 Conclusions and Future Research

Applications with immersive audio–haptic support can be developed for the blind, based on 5.1 audio and a force-feedback joystick using DirectX as the hardware abstraction layer. In these environments, care must be taken to allow users to identify the position and movement of sound sources. High frequency sounds, a minimum player turning angle of 45º, ample character motion and a large soundscape will help to create a near-realistic and enjoyable application.

More research is needed to create a better guide. Two new games ("Jedi Trainer" and "Dancing Mat") with automatic collection of user reactions will be under final user testing in 2006. A new haptic editor with Lego tiles will allow blind users to create their own stories and soundscapes, based on previous work experiences [2], [16]. This editor is related to SHADE and will become available for testing in 2006. We expect to collect a set of rules for future use in the industry, in the shape of a job skills guide and a skill level qualifier for simulation and games. Our metrics would relate to the variation of sound length, intensity and frequency; sound simultaneity and sound movement, with and without haptic reinforcement.

We expect these guidelines to allow less experienced developers to produce predictable and enjoyable applications for blind and sighted people.

References

1. Ortega, P.: Juegos Educativos Para Ciegos: El Ahorcado. Final-year project, Facultad de Informática, Universidad Politécnica de Madrid (2004)
2. Sánchez, J., Lumbreras, M.: Ambientes Virtuales Interactivos para Niños Ciegos, Department of Computer Science, University of Chile (2002)
3. Sánchez, J., Sáenz, M.: 3D Sound Interactive Environments for Problem Solving, Department of Computer Science, University of Chile, ASSETS'05 (2005)
4. Kurniawan, S. H., Sporka, A., Nemec, V., Slavik, P.: Design and User Evaluation of a Spatial Audio System for Blind Users, Department of Computation, UMIST, Department of Computer Science, Czech Technical University in Prague, 2004 ICDVRAT (2004)
5. Microsoft Corp, DirectX Tutorial, `http://msdn.microsoft.com/archive/default.asp?url=/archive/en-us/directx9_m/directx/directx9m.asp` (2005)
6. Microsoft Corp, DirectSound Reference Manual `http://msdn.microsoft.com/library/ default.asp?url=/library/en-us/wcedsoun/html/_cerefDirectSoundReference.asp` (2005)
7. Klatzky, R. L.: Procedures for haptic object exploration vs. manipulation. In: Goodale, M. (ed.): Vision and action: The control of grasping. Ablex, New Jersey (1990) 110-127
8. Babuglia, M., Cuiñas, G., Fernández, N., Lazar, M., Pereira, P.: "Audiojuegos, SHADE: Simple Haptic Acoustic Development Engine", Systems Engineering Degree, Facultad de Ingeniería, Universidad ORT, Uruguay (2005)
9. Algazi, V., Duda, R., Duralswami, R., Gumerov, N., Tang, Z.: Approximating the head-related transfer function using simple geometric models of the head and torso. J. Acoust. Soc. Am. 112 (2002) 2053-2064
10. Alonso, F., Fuertes, J., Martínez-Normand, L., Szabo, H.: Análisis de Percepción en Entornos Acústicos Envolventes con Asistencia Háptica para Ciegos. AST 2005, Rosario, Argentina (2005)
11. Algazi, R., Duda, R., Thompson, D., Avendano, C.: The CIPIC HRTF database. WASSAP '01, New York (2001)
12. Blauert, J.: Spatial Hearing: The Psychophysics of Human Sound Localization. The MIT Press Revised Edn (1996) 137-150, 237-246
13. Immersion Studio for Gaming, `http://www.haptech.com/developer/technology/tools/index.php` (2005)
14. Magnusson, D., Rassmus-Gröhn, K., Sjöström, C., Danielsson, H.: Haptic 3D Object Recognition – a Study with Blind Users, Vision 2002, Göteborg, Sweden (2002)
15. Sato, M.: Development of String-based Force Display: Development of String-based Force Display, SPIDAR Precision and Intelligence Laboratory, Tokyo Institute of Technology (2002)
16. Bussell, L.: Touch Tiles: Elementary Geometry Software with a Haptic and Auditory Interface for Visually Impaired Children, EuroHaptics 2003 (2003)
17. Dolby Laboratories Inc.: Room Layout and Speaker Setup, `http://www.dolby.com/consumer/home_entertainment/roomlayout.html` (2005)

The Amodal Communication System Through an Extended Directional Input

Georgios Yfantidis and Grigori Evreinov

Dept. of Computer Sciences, University of Tampere
georgios.yfantidis@nokia.com, grse@cs.uta.fi

Abstract. Multi-modal interfaces have been overflowing HCI research, incorporating the different senses, to provide adequate feedback or input for human-device interaction. The plethora of sensory combinations that this "creeping multimodalism" implies seems to be creating an oxymoron when it is used as a solution to help people with sensory problems and/or limitations dealing with interfaces. A better solution for those people would be to use systems where the traditional senses are obsolete as driving factors of the interaction, and they are only used as peripheral aids. The quest for such an amodal user experience is the object of our current research.

1 Introduction

The directional gesture system is an example of potential amodal interaction techniques when combined with touchscreen input. Blind people could use such a system as a text entry tool on touchscreen based mobile devices and PDA's [4, 5, 9, 10]. The subject of our current research is to employ gestures that would be easily associated with integrative notions [2, 6, 7]. The goal is to maximize the inherent potential for intuitiveness that lies in touchscreen based input [1] employing directional gestures, and to ultimately provide blind users with an amodal input tool, that could be customized to be used in different applications [3, 8].

We are focused towards advanced systems that could adapt to user performance and interpret the user input and intention. Systems like that can be powerful and useful when there is an acquired or intuitive appropriation between direction and integrative notion that creates a symbolic/generic language inherent to human brain. The different directional gestures could, for instance, correspond to notions and categories, which are specific for mathematics and represent a set of particular actions which may be performed with objects, numbers (e.g., Chinese characters for numbers) or graphs.

Addition, subtraction, multiplication, and division (fraction expressions, numerators and denominators) could be associated with directional gestures. For example a rightward gesture could correspond to the notion of addition "+", while a leftward gesture could be subtraction "-". The same appropriation could be used to define numerical difference and equality (left-side/leftward and right-side/rightward), geometrical relationships of parallelism such as orthogonality, similarity, congruence, tangency and so on. In a more basic approach eight different directions could be associated with eight numbers or different alphabet characters [10].

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 1079 – 1086, 2006.

The system that we are developing is using an input of straight-line gestures that are classified in the four basic and intermediate directions. Furthermore, it is able to discern between parallel e.g. "// " and sequential "– –" gestures. This would already enable a wide range of gestures that could be tied to a "vocabulary" of corresponding inherent notions. In order to determine the factors that differentiate the two-part gestures into single and double we needed to test a vast number of handwriting patterns to build a reliable recognition algorithm for user input. This test was the starting point of our research and helped us build amodal communication system.

2 Method Design

2.1 What Makes the Directional Gestures Differ

The amodal system that we plan to construct will have to be able to recognize three types of gestures. The main challenge was for the system to recognize if a gesture is single, or the first part of a double gesture. Ten subjects from the local university were recruited for preliminary experiment to gather a base of three types of directional gestures. The goal of the experiment was to find the parameters (predictor-factors) defining when recognition and classification can be performed regarding uncompleted two-part directional patterns.

Setup. The hardware used in the experimental set-up was a pen tablet Wacom Graphire4. The software recognizes and stores data for two-part gestures done in the Wacom tablet, using stylus. Each of the two-part gestures comprised of two vectors, the first gesture, the second gesture and the distance between their starting points. The parameters that were collected were the length (R1, R2, R3-distance) and the time (T1, T2, T3) accordingly. Those parameters are redundant to parse the data input stream. However, for directional input techniques "interrupted sequences" of handwriting patterns were never used before because only uninterrupted gestures seemed to lack discrimination problems [5].

The subjects were instructed on the concept of single and double gestures. They were told that the system receives an input of three types of gestures, and thosecould be

- two single gestures implemented sequentially with a little pause as naturally as possible,
- one double gesture composed of two parallel handwriting strokes and
- one double gesture composed of two sequential handwriting strokes.

The task that was given to them was to perform a number of gestures using pen input on a tablet. The gestures had to be straight lines toward specific directions.

Each subject had to complete three sessions of tests. Each subject would have to produce gestures following all four of the basic directions (up, down, left and right). The purpose was to gather data for the four basic directions, at least, not to compare them and focus on possible differences as such, but in order to observe if a common pattern would arise.

The first session was a collection for single gestures. The subjects were told that they had to produce 250 two-part gestures (where each part is a "single gesture").

The second session was about making 250 double sequential gestures "– –" where the two parts had to be short enough to form together a complete pattern.

The third test session exactly the same with the second except that 250 double parallel gestures "// " had to be produced instead.

Sound signals were used to give the subjects feedback that the recording of the input pattern was completed and the number of gestures was displayed. At this phase the system did not make any comprehensive analysis of input data, but was only recording the number of Down-Up events and time per stroke. When the conditional task was inadequately matching the template, the data were not adopted and an error sound was given.

Results of the Preliminary Test and a Pilot Model

In total 30000 (10subjects × 250 trials × 4 directions × 3types) handwriting gestures were collected. After processing the data of the preliminary test we found out that the defining factors for the different types of two-part gestures (single or double) were the length (R1) of the first vector (which seemed to be longer for single gestures), and the time (T3) of the intervening space between the gesture vectors in a sequence.

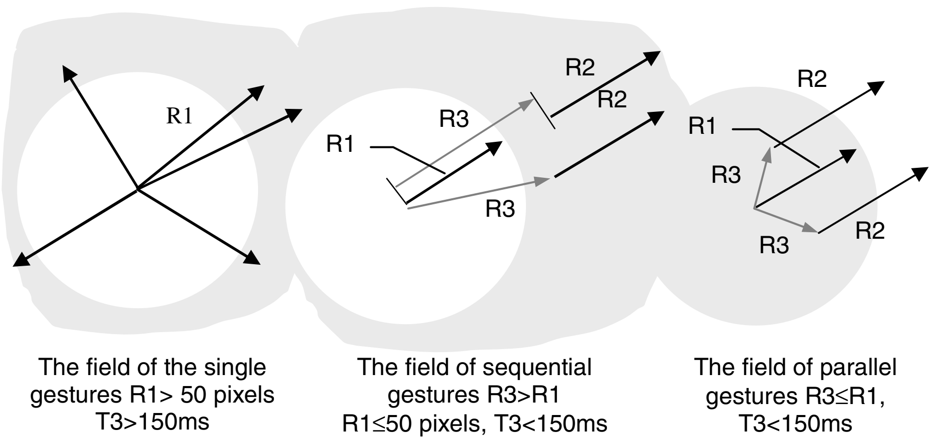


Fig. 1. Handwriting directional gestures from the arbitrary location with touch input

Besides that, T3 can efficiently split or integrate single and double gestures when the length R1 is ambiguous. R3 is the second criteria to differentiate parallel and sequential gestures (Fig.1).

2.2 Auditory Visualization of the Temporal Scale

At first, we eliminated the navigational problems by making the interaction independent of specific starting location in the tablet [10]. When the main discrimination factors were identified we built a new feature in the testing software, which could guide/synchronize the blind user actions (handwriting strokes) to augment input space and, finally, increase reliability of interaction technique. We supposed that the length of strokes could have less variation when the "temporal window" has a "visible" temporal scale. In particular, an auditory strobe-like signal was integrated in order to provide pace in blind handwriting. The maximum time that we assigned for the production of directional gestures in the experiment was 2000 ms, which could be interrupted when a two-part gesture was considered as completed. We added a sound signal that would mark the interval limit between two single gestures repeated every 400ms until the 2000 ms "deadline". Those were not absolute guides on how to employ pace, but the subjects (of the algorithm verification test that followed, and is described in section 3) were informed that they could follow the signal to ensure completion of the handwriting strokes within the appropriate time limits.

2.3 Description of the Algorithm

When the pen touches the tablet for the first time (MouseDown event 1) the system saves the position X_0Y_0 and the time T_0. When the pen is lifted (MouseUp event 1) the new data $X_1Y_1T_1$ are also saved. At this point, the timer for the 150ms interval starts up. Also the length R1 and angle 1 are calculated.

When the length of vector R1 exceeds 50 pixels then the first assumption is made regarding a "single type" gesture. The next assumption will follow when the 150 ms have expired. If so, then the timer is stopped and the gesture is recorded as a "single type". Nevertheless, if the pen touches the tablet for a second time (MouseDown event 2) then the properties $X_2Y_2T_2$ are saved. If 150 ms have not yet reached their expiration (the timer is still active) then the timer is stopped at that point and the gesture can be considered as a "double type" stroke.

The length between MouseDown event 1 and MouseDown event 2 is recorded as R3. The data $X_3Y_3T_3$ are also recorded for MouseUp event 2. Herewith, when R3 is longer than R1 then the gesture can be classified as a "double sequential" type. If R3 is shorter than or equal to R1 then the double gesture is classified and recorded as a "double parallel" type.

The length between MouseDown event 2 and MouseUp event 2 is saved as a vector R2 and the angle of this segment is also calculated as angle 2. Those values refer to the second segment of the double gesture.

Certainly, in handwriting stream, the strokes with non-equal angles might be referring to different patterns. Therefore, the final check for two-part gestures has to do with comparing the angles, to verify if both segments are pointing to the same direction. When angle 1 equals angle 2 then the gesture is valid as a double type, and if not the system rejects the gesture as a non-recognizable type (non-commensurable to template) that is not used in the test.

3 An Extended Directional Input Evaluation

3.1 Verification of the Recognition Algorithm Setup

The next step was to test the recognition algorithm. A new experiment was organized having as a main purpose the verification of the reliability of blind input and recognition of extended directional gestures. The test was carried out with 8 subjects. The hardware setup was the same with the preliminary experiment where the subjects just performed particular gestures.

For the extended test, we categorized the gestures in directions and types, and the new experimental software expected a specific gesture as input each time. When the performed handwriting pattern matched the expected one, the gesture was recorded as valid; otherwise it was categorized as an error.

The system had the capability of recognizing 24 different handwriting patterns all of which were exclusively comprised of directional strokes. The three main categories were single gestures made of one vector, and then there were double sequential and double parallel gestures. Each of those types could be performed and recognized in eight different directions. The directions were labelled (1-8) and they were situated between the following degrees (337-23-68-113-157-203-247-293-337). In other words, these were the basic directions up, down, left and right, and their in-between directions (e.g., left up).

All in all, the required gestures for each test session were 240 (24 gestures repeated 10 times each). The gestures were not expected by the system in a linear way, but they were required in a random order. When 10 gestures of a specific type were made, this type was not requested any more within this test session. Each subject had to complete 8 sessions.

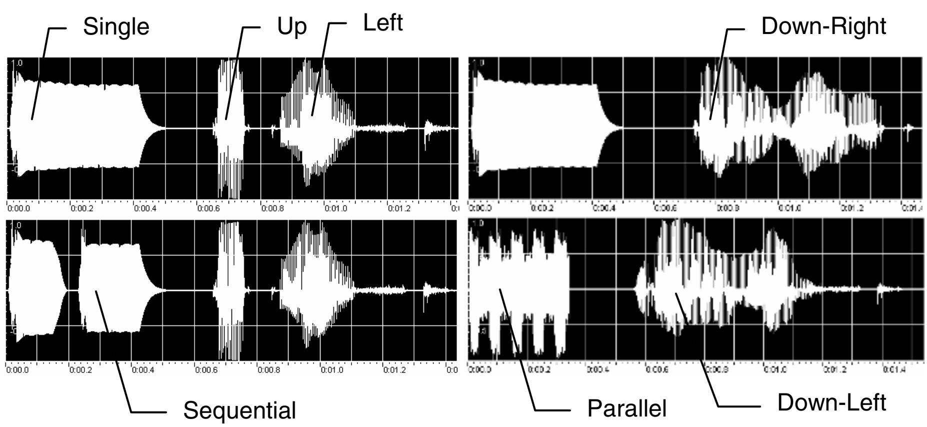


Fig. 2. Non-speech cues to assign a type of gestures, and speech cues for direction in testing

One gesture had to be entered at the time after the "requirement" was given by the system. The required gestures for the test had to be made in blind mode. There were only auditory cues as to what kind of gesture was expected each time. The cue could be repeated at will by pressing a button, on subject request. A combination of speech and non-speech cues were used in order to convey the direction and the type of gesture that was required. The type of gesture single, sequential or parallel was described by non-speech cues and the actual directions were given by speech cues. (Fig.2)

3.2 Results and Discussion

R1 is the principal discriminating factor for single and double gestures and this is illustrated in Fig. 3 where the values for different types of gestures are clearly over and below the discriminatory limit of 50 pixels, allowing a "comfortable" classification of gestures, which is crucial for the recognition system reliability.

T2 stayed well below the maximum limit of 150 ms (Fig. 4) giving comfortable recognition possibilities for the secondary discriminatory factor for single and double gestures.

Errors were in very low levels for plain single gestures, while for double gestures the sequential ones proved to be the more error prone. The mean errors per direction demonstrated that even for beginners, errors for the most "difficult" type of gestures (double sequential) were between 7 and 18%. We expect the errors to decrease significantly as the subjects will gain more experience in producing the gestures.

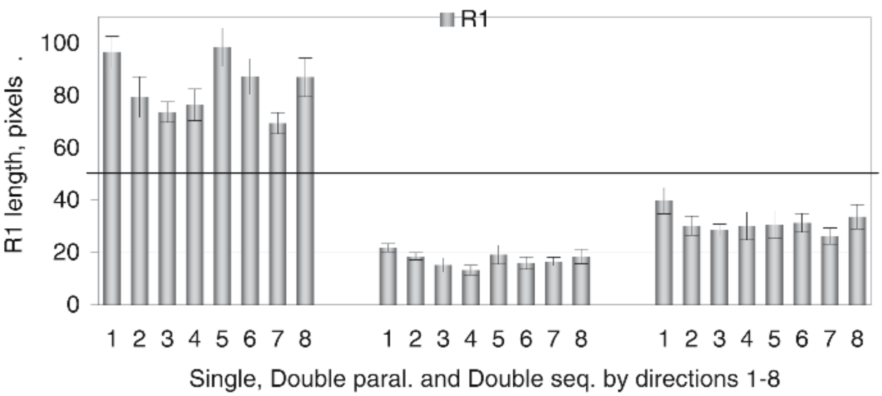


Fig. 3. Mean R1 for single, double parallel and double sequential gestures. The black horizontal line shows the limit for differentiation of recognition.

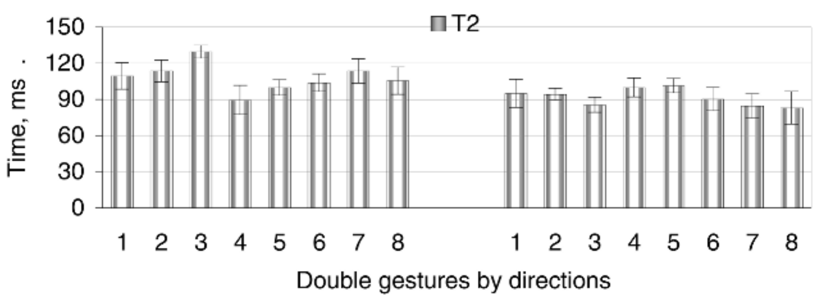


Fig. 4. Mean T2 for double parallel and double sequential gestures

The values that were received for R1 and R3 follow the algorithm. Fig. 5 illustrates that the length variations for R1 and R3 are directly linked to the gesture type in all directions.

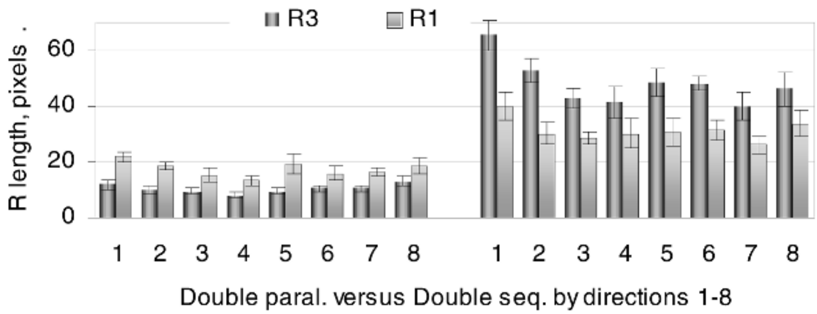


Fig. 5. Mean values for R3 and R1 for double gestures by direction and type of gesture

3.3 Conclusions

Analysis of the extended test results verified the functionality of the algorithm and the correctness of our model. We can indeed differentiate between single and double, parallel and sequential gestures produced in blind mode based on values of length, time, and angles.

The sound strobe signal was a valuable confirmatory feedback for the user, which helps her/him produce reliable single or two-part gestures of different properties (single-double, eight directions). From the system point of view the user-input integration by pace (synchronization of handwriting behaviour) decreased the variation of the parameters used for gesture recognition more than twice.

4 Future Research

Based on the encouraging results of the experiments, we can now focus our research on applications of the extended directional input technique. We were initially oriented on building a system that would use the amodal patterns for human-device integration. Such a communication system can be based on a simple gestural input accessible to blind people. The crucial step now is to carefully construct a hierarchy of the different types of directional handwriting patterns that now became available, with appropriate system functionality.

The gesture driven amodal system is versatile enough to encompass different functions varying from text input, to arithmetic expressions or even manipulation within an interface (such as navigation). Furthermore we plan to expand the assistive secondary feedback that can be used as general guidance or confirmation for the blind users. We plan to integrate tactile feedback through vibration to supplement the current sound signals. This will result into a system that will use amodal input and will in addition encompass different types of secondary feedback to enhance the user experience, and assist the blind user in more customized and effective ways.

Acknowledgments

This work was financially supported by the project “Multimodal Collaboration Environment for Inclusion of Visually Impaired Children” funded by the EU Commission, IST-2003-511592.

References

1. Arato, A., Juhasz, Z., Blenkhorn P., Evans, G., Evreinov, G. Java-powered Braille Slate Talker. In Proc. of the 9th Int. Conference on Computers Helping People with Special Needs. ICCHP 2004, LNCS Vol. 3118, Springer-Verlag, 2004, pp. 506-513.
2. Greco, A., Caneva, Cl. From Action to Symbols and Back: Are There Action Symbol Systems? In Proc. of XXVII Annual Conference of the Cognitive Science Society, 2005. Website, 2005: http://www.disa.unige.it/DOCS/PERS/greco/pubblicazioni/stresa.pdf .
3. Himberg, J., Häkkilä, J., Kangas, P., Mäntyjärvi, J. On-line Personalization of a Touch Screen Based Keyboard. In Proc. of the 2003 Int. Conf. on Intelligent User Interfaces (IUI 2003), ACM Press, 2003, pp. 77-84.
4. Kurtenbach, G. and Buxton, W. User learning and performance with marking menus. Proc. of CHI '94, (1994) 258-264. At: http://www.billbuxton.com/MMUserLearn.html
5. Isokoski, P. A minimal device-independent text input method. Report A-1999-14, Department of Computer Science, FIN-33014, University of Tampere, Finland, 1999. Available at: http://www.cs.uta.fi/reports/pdf/A-1999-14.pdf .
6. Ventrella, J.J. “Yaw, Pitch, Roll <> No, Yes, Maybe”, Website, 2005 http://www.ventrella.com/Ideas/YawPitchRoll/yaw_pitch_roll.html .
7. Wobbrock, J.O. The Benefits of Physical Edges in Gesture-Making: Empirical Support for an Edge-Based Unistroke Alphabet. Extended Abstracts of the ACM Conf. on Human Factors in Computing Systems (CHI '03). ACM Press, pp. 942-943.
8. Wobbrock, J.O., Myers, B.A. and Kembel, J.A. EdgeWrite: A stylus-based text entry method designed for high accuracy and stability of motion. In Proc. of the ACM Symposium on User Interface Software and Technology (UIST '03) Vancouver, British Columbia November 2-5, 2003, ACM Press, 2003, 61-70.
9. Yfantidis, G., Evreinov, G. (2005) Blind Text Entry for Mobile Devices. In Proc. of AAATE 2005, Lille. France, IOS Press, pp. 246-250.
10. Yfantidis, G., Evreinov, G. Adaptive Blind Interaction Technique for Touchscreens. J. Universal Access in Information Society, Springer, 2004, vol. 4(4), available at http://dx.doi.org/10.1007/s10209-004-0109-7.

Multimedia Browser for Internet Online Daisy Books

Piotr Brzoza[1] and Dominik Spinczyk[2]

[1] Silesian University of Technology, Computer Science Department, Akademicka 16
44-100 Gliwice, Poland
piotrbr@polsl.pl
http://www.key.org.pl
[2] Silesian University of Technology, Biomedical Department, Akademicka 16
44-100 Gliwice, Poland
dspinczyk@polsl.pl

Abstract. Visually impaired people have limited access to printed media: books, magazines, newspapers etc. Only one to five percentage of printed books and magazines are published in the form accessible for this group of readers. Nowadays printed materials are prepared on computers in digital form. These documents can be accessible for all readers. Paper describes a computer system enabling interactive online presentation of multimedia Daisy books over the Internet. The system cooperates with the Internet multimedia library computer management system. The main goal of both projects and their execution, is easy and effective access to information for visually impaired people. We focus on new feature of our DaisyReader which allows interactive voice reading of math formulas.

1 Introduction

Visually impaired people have limited access to printed media: books, magazines, newspapers. Only a small percentage of printed books and magazines are published in Braille or are recorded on cassettes or CDs. The cost associated with those publications is very high. It limits and delays the amount of available publications.

Computers equipped with screen magnifier, screen reader or talking multimedia browser enable blind users unlimited access to information. Internet library portals, friendly to visually impaired readers allow for easy and effective access to catalogues of public and academic libraries. The next step in accessibility of printed materials is their structuring and change to digital form. The main goal of current projects is to perform such structuring, which will enable presentation of the book's content in multimedia form adapted to the needs of differently handicapped reader groups.

2 State of Art

2.1 Accessibility of Library Web Services

Accessibility and usability tests were conducted to assess to public and academic library portals in Poland. Library web pages were evaluated using Bobby service

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 1087 – 1093, 2006.

against W3C accessibility standards. Blind people using assistive software performed developed set of test tasks. A number of accessibility issues were detected: difficult navigation, graphical links without alternative text description, untitled frames and inconsistent labelling for form elements. The library systems were accessible to experienced blind computer users but sometimes difficult in use.

2.2 Digital Publications Accessibility

Not long ago visually impaired people had access only to books in Braille format and audio books recorded on cassettes or CD. Affordable personal computers allow access to printed publications by new innovative ways. An example of computer access to information is the reading of printed text using a scanner, optical character recognition software (OCR) and a program that change text to speech (TTS).

Today many publishers offer both printed and electronic book versions (called e-Books). E-Books can be read on e-Book reading devices, personal computers, and palm-sized personal digital assistants. Electronic books are published in several e-Book formats. To read them users need various e-Book reading software.

Digital Rights Management (DRM) used by e-book reading software causes additional accessibility problems. The e-Book reading software presents the text in graphic form that does not allow the text to be transferred to the clipboard. This pre vents the

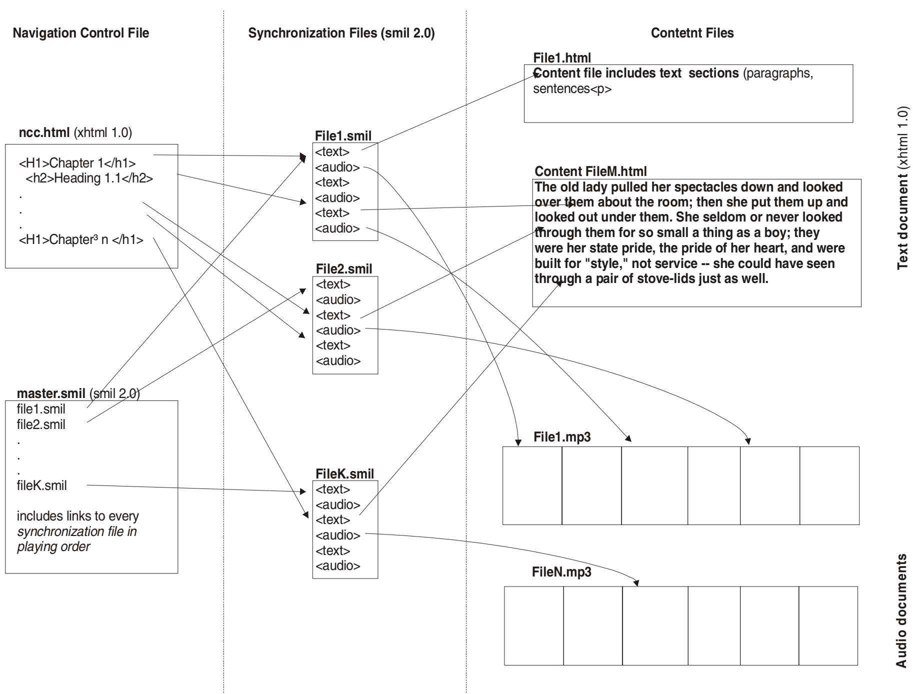


Fig. 1. Daisy book structure

duplication of the book content. Unfortunately it also prohibits the screen reader from accessing the e-book's synthetic speech feature. The end result is that blind people are denied access to the e-Book, because the screen reader does not work correctly.

The screen reader software they use intercepts the information being displayed on the monitor and uses a speech synthesiser to read it aloud.

As a solution to this problem the companies such as MICROSOFT and ADOBE are proposing to extend functionality of e-book readers with an additional audio interfaces that would allow e-books to be accessed by people who are visually impaired.

2.3 Daisy Standard

The Daisy Consortium founded in 1996 representing publishers and libraries for the blind has developed a new digital multimedia book standard, accessible to visually impaired readers (DAISY 3.0 ANSI/NISO Z39.86-2002). Daisy books present book content in multimode form including text, audio, and graphics. Readers can easily navigate in logical book structure by: chapters, headings, pages, paragraphs and sentences. Main daisy book structure is presented in figure 1. Daisy books can be played using hardware or software players. Currently world wide there are about 130000 available book titles.

3 System Deployment

We present interactive system for online multimedia daisy books presentation, which main components are presented in figure 2.

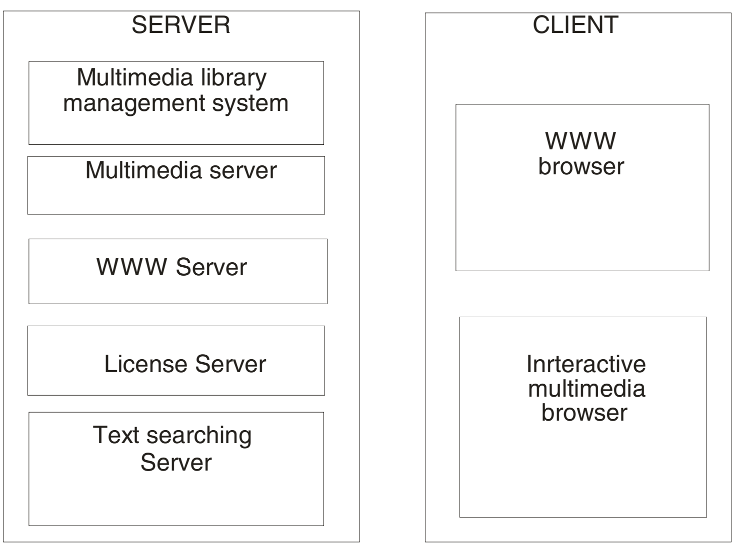


Fig. 2. Multimedia system components

3.1 Management System for Internet Multimedia Library

We present a system, which manages the virtual library of multimedia publications. The system is a result of a research project conducted by The Silesian University of Technology together with the Academic Library and School for the Blind in Poland. The design of the system allows:

- cataloguing and collecting of multimedia publications like e-books, e-magazines, digital talking books, digital music and movies
- assures secure Internet access to the library resources by registered users
- provides management of reader's orders
- distributing of the ordered publications on CDROM disks

The presented library system is running in Linux environment on Pentium multiprocessor servers. The system is managed with the web user interface and standard Internet web browsers. The library is easily accessible to visually impaired people. The user web interface was designed with the special attention to requirements of blind and low vision internet users and allows for direct access to information on internet pages, easy navigation and adjustment of font size, color and contrast as specified by the individual user. WWW library service was build according to the W3C-WAI web content accessibility guidelines. Accessibility of the library pages was tested with the software for visually impaired users. Zoomtext, Jaws and Window-Eyes were used during testing.

We present system services as data flow diagram in figure 3.

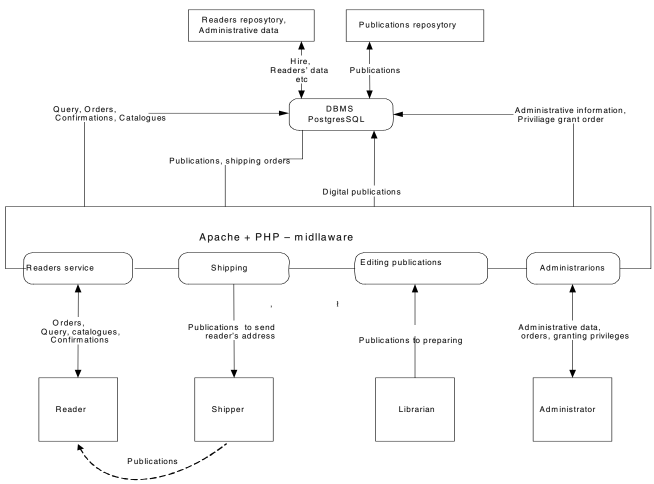


Fig. 3. Data flow diagram of management system

The presented library system offers the following services: all Internet users can search the library catalogues and browse detailed search results with information about found publications. This allows potential readers to get acquainted with offered publications and the user interface of the library system. All other services are available only to registered users. This guarantees secure and controlled access to offered publications. Users of the system are authorized during login with securely passed user identifier and password. Handicapped readers using standard web browsers and assistive software can search publications in the catalogue by title and author. Found publications are presented as a list. This allows handicapped readers to easily browse search results. After the publication title link is clicked on, the page with the detailed information about the publication is presented to the user. Readers can download publications to their local computer disk, browse publications online, add publications to the cart and browse its contents as in Internet shops. Publications collected in the cart can be ordered on the CD ROM disks. This allows for the alternative way of delivering digital talking books to readers who do not have fast Internet access. Readers can also browse the history of borrowed publications and suggest new publications for the library. The library system offers its users context sensitive help available in the navigation bar. The font size and color used on the library pages can be adjusted on a per user basis according to the information stored in the personal user profile.

The system offers the following services to library personnel. Librarians can browse and modify the list of new publications proposed by readers. They can choose which publications to be prepared in accessible format and input them into the library server system. The system manages and allows for a number of librarians to work independently on preparing their publications. Next using the web browser he or she enters the catalogues data and publication file into the library server. After the data is verified, the publication is available to readers. The library system enables the efficient management of publications, reader accounts, orders, and generation of statistical library usage reports.

3.2 Interactive Multimedia Daisy Books Browser

Currently available daisy players allow reading multimedia Daisy books to be stored on compact discs (CD) or local computer hard discs. Books recorded on CD-ROM's are collected personally or ordered by mail in the library. Some libraries offer digital books by Internet. This method requires full book contents download over the Internet. Books in Daisy format range from several to hundreds of megabytes in size. Downloading large amounts of data makes this approach to books' distribution both difficult and time consuming. Access to information contained in multiple books is very difficult and multiple books information searching is practically impossible. In our continuous research and development of the multimedia library system we have designed and developed new multimedia Daisy book browser. The new software Daisy reader allows playing Daisy books online over the Internet or in the standard way, from CD or from the local hard disk. Online Daisy books are played from a multimedia server and are available for reading after a few seconds from being found in the library system. Books audio, text and graphics content are presented synchronously. Consecutive book pages are displayed. Sentences are highlighted with simultaneous audio being played. The book's index allows access to selected chapter. The reader can also navigate through the logical book structure by: chapters, headings, pages, tables, paragraphs and sentences.

Similar as in printed books readers can add bookmarks with text, audio notes and exchange bookmarks list each other. Browser offers searching text function which allows playing narrator's speech and presenting highlighted text from place where it has been founded.

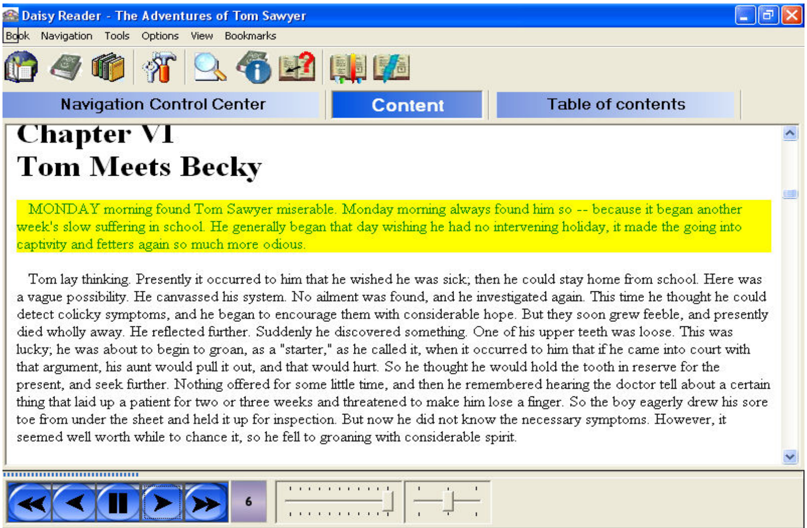


Fig. 4. Main window browser

Browser user interface is customizable to different users group: blind, low vision, dyslectic, mobile impairment. Blind people can use browser with assistive software or use it in self voicing mode. We extend browser functionality which allows read aloud text DAISY book with synthetic voice. New navigation commands enable readers interactive audio browsing: sentence reading, word spelling and reading structural information like tables, math formulas. Interactive audio presentation of math formulas and tables allows preparing advanced DAISY scientific, technical books and educational materials for students, mathematicians and scientists. Our research, which cooperates with DAISY MathML Project, lead to extends DAISY standard.

3.3 System Implementation

Our interactive multimedia Daisy book browser works together with multimedia Helix Universal server. Multimedia Helix server is integrated with the multimedia digital library management system. Using the web browser, users can search and browse library catalogue, after Daisy book selection system generates an encrypted license file (file with extension .dtb) this file contains access rights to the selected book. The web browser automatically starts the interactive multimedia Daisy book browser with the selected file. Next Daisy book browser establishes a connection with the streaming Helix Universal server. Text and graphics files and meta data describing book structure are accessed from WWW server. Multimedia book content in DAISY format is divided into parts and stored in many files: text xml, audio mp3 and graphics files. This is essential to continuous online book presentation over the Internet. Additional browser mechanisms preload book fragments and allow to present book contents without interruptions. Files are accessed in parallel with audio multimedia stream. Loaded xml files (containing book text) are buffered in Daisy browser memory which allows for smooth navigation. The user can search for information in the document before the document is loaded. The text search server implements this function. After the search request is processed and information about found text is passed back to browser. This information allows Daisy browser to playback book contents from any fragment that meets the search criteria.

3.4 System Scalability

Selected multimedia streaming server (Helix Universal Server) works in a multiprocessor environment and allows for simultaneous data transmission and multi user service. During our work we tested the system's scalability with a dual processor Pentium IV server running Linux operating system. Scalability tests were conducted using special client and server applications with client applications count from 2 to 32. Test applications working together with an independent coordinating server allowed collecting various transmission quality statistics.

4 System Deployment

Presented system was installed at library for the blind in Poland. It successfully serves thousands users. Currently the library offers 2000 publications in the polish language. Visually impaired computer users in Poland have gained easy access to growing collection of books and magazines in digital format.

The library system will also be installed in computer centre for blind and low vision students in Warsaw University, Poland. Adaptation centre of the University is preparing educational materials, scientific and technical books in Braille and other accessible digital formats. The library system will make possible cataloguing and collecting accessible educational materials prepared by other universities adaptation centers. The library will provide the requested publications to the registered disabled students from several universities. The preparation of educational materials in a group of universities can be less redundant and it is possible to produce more books in accessible formats.

Our DaisyReader can be free downloaded from `www.key.org.pl` . For more information visit `www.iddw.org` .

References

1. Brzoza P., Spinczyk D.: Online Internet Interactive System for Multimedia Daisy Books Presentation. Accessible Design in the Digital World Dundee (2005)
2. Brzoza P., Moroz P.: Internet Multimedia Library Accessible to Impaired Readers. ICEVI European Conference, Chemnitz (2005)
3. Brzoza P.: Presenting structural information in multimedia documents. Computer Networks Conference, Zakopane (2005)
4. Brzoza P.: Presenting accessible information in multimedia environments. Workshop New strategies for accessible information provision. Paris (2004)
5. Brzoza P.: Virtual multimedia library accessible to blind people. Technology And Persons With Disabilities, Los Angeles, (2003)

Automated Book Reader for Persons with Blindness

Malek Adjouadi, Eddy Ruiz, and Lu Wang

Center for Advanced Technology and Education, Electrical & Computer Engineering
Florida International University, 10555 W. Flagler Street, Miami FL 33174, U.S.A.
adjouadi@fiu.edu, eddyruiz33177@yahoo.com, lwang001@fiu.edu

Abstract. This research introduces a new automatic book reader for persons with blindness. The objective is to design a fully integrated system that is relatively fast and yet inexpensive and effective with a high reading accuracy. Through the use of two inexpensive light weight cameras, a book holder and using regular lighting of an office or a lab environment, this integrated system addresses through software development (a) the mathematical foundation of perspective distortion introduced by page curvature of an open book, and of barrel correction introduced by the inherent nature of image capture of the camera; (b) image preprocessing for finding lines and characters in a given image; and (c) the implementation of a fast neural network that takes as input the findings of step (b) and provides an audible read out through a speech synthesis engine.

1 Introduction

The focus of this study is on the creation of an interface that will provide an audible readout of a digitized document as an assistive technology interface for persons with visually impairment and blindness. The interface requires the acquisition of a digital representation of the document, processing of the document in order to extract the readable characters, and classification of the characters so that they can be read aloud. Image acquisition must be reliable and tolerable of document placement errors that could affect the recognition of characters within the document. This acquisition process is further supported through automated corrections based on barrel and perspective distortions that open books introduce.

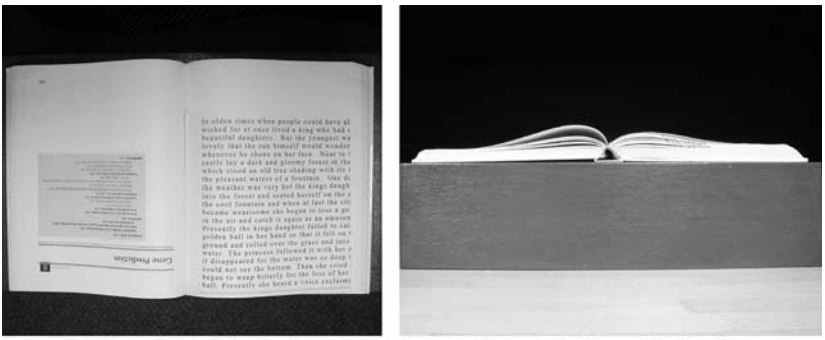
The literature shows extensive work in the area of character recognition, but current methodologies are either inadequate or constrained during document acquisition. This interface extends beyond problems of skewness introduced by document readers that use flat-bed scanner [1,2] and augments the capabilities of the traditional Optical Character Recognition (OCR) systems [3,4]. The goal of the interface is to automate and simplify the task of reading a book while providing a user-friendly environment. The interface also responds to the main concerns of (a) providing a method of image acquisition that maintains the integrity of the source (b) overcoming character recognition errors created by document placement errors and other imaging issues such as the perspective effect, the warped effects of open books due to page curvature, and (c) determining a suitable classification method for accurate character recognition by proposing a new neural network algorithm. The challenge therefore remains in the ability to read from any open book by developing the mathematics to de-warp the curvature of pages, a problem that is certainly not

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 1094 – 1101, 2006.

encountered with flat documents, as well as alleviating barrel distortions and noise problems. The main goal is to extend access to books beyond integration of e-books [5] and beyond information access through computers [6-8].

2 System Design

Equipment and materials used to design the system include: (1) Two Cameras: one lateral and one above the book, both are Sony DSC M1 digital cameras; (2) Book support: A simple wooden box with straight lines and flat surface, used as means to facilitate image capture of the book as shown in Fig. 1; and (3) Background paper material or cloth with solid color, preferably black, used to facilitate extraction of the book curvature through the lateral camera.



(a) Top view of the book (b) lateral view of the book

Fig. 1. Image Acquisition Process

3 Mathematics for Image Acquisition and Removal of Distortion

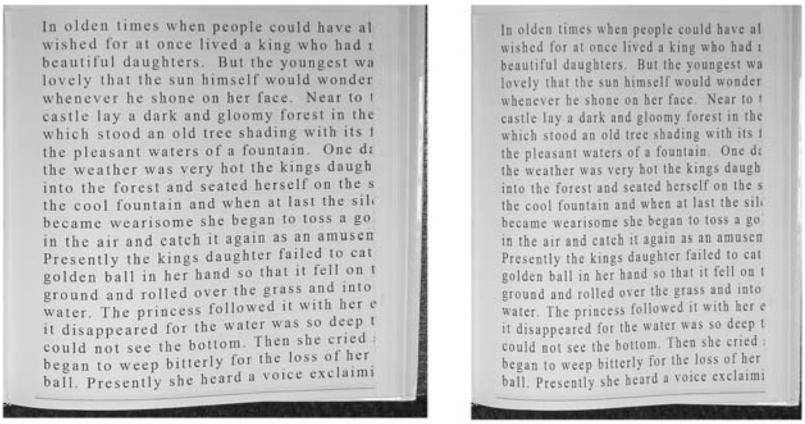
The support box the book is placed upon is used as a reference frame from which a lateral image and top image of the book are extracted. For processing, we need to find out the same starting and ending points of the two images and then find out the corresponding points on each image. This is accomplished using simple convolutions of first derivatives to the images and the two boundaries of the two images are co-registered. Page curvature is assessed through edge detection performed using the Laplacian of Gaussian operator on the lateral image shown in Fig. 1(b). The distance between the camera and the book decreases with increased curvature.

3.1 Barrel Distortion Correction

The correcting function for barrel distortion is a fourth order polynomial:

$$r_{src} = a \cdot r_{des}^{\ 4} + b \cdot r_{des}^{\ 3} + c \cdot r_{des}^{\ 2} + d \cdot r_{des} \tag{1}$$

Where r_{src} is the distance between the pixel and the center point of the original image, and r_{des} is the distance between the pixel and center point in the corrected image. The distance is normalized with respect to the width of the image. We solve the polynomial equation and take the root as r_{des}, which is closest to r_{src}. Based on the camera used, in this case the SONY DSC M, the solution to (1) is [a= 0.016048,b=-0.042510,c=0 and d=1.026462]. Note that these values are computed only once depending on the camera used. We can see in Fig. 2 that the boundary is corrected, which means that the pixels far away from the center are projected in the more distant points. That is exactly what the barrel correction is supposed to do [9,10].



(a) Original Image (b) Image after Barrel Correction

Fig. 2. Barrel Distortion Correction

3.2 Perspective Transformation

Using the perspective transformation between any two points $P_i(X_i, Y, Z_i)$ and $P_j(X_j, Y, Z_j)$ in the real world and their projection in the image space (x, y), with Z being a reference height, a weight w is established $w_k = (Z_k - f)/(Z - f)$, where k=[1,2, …, N], with N being the number of points in the horizontal axis, and Z_k is the distance between the kth pixel and the top camera. So, for any x_i, x_j, y_i, y_j,

$$d_{xji} = x_j \cdot w_j - x_i \cdot w_i = f \cdot \frac{X_j - X_i}{Z - f}; \; d_{yji} = y_j \cdot w_j - y_i \cdot w_i = f \cdot \frac{Y - Y}{Z - f} \tag{2}$$

Where d_{xji} and $d_{yji} = 0$ are the distance between i and j pixels in x and y directions, respectively. After correction, the two points will be on the same line in the image. Through equation (2), the distortion in the Y direction is recovered (see Fig. 3(b)).

3.3 Character Extension for Text Projection to Overcome Curvature

In this step we need to eliminate the distortion in the X direction. Let us denote by X_i^r the real position of point i in the X direction. Then we set a factor

$s_{ji} = (X_j - X_i)\big/(X_j^r - X_i^r)$. By using Equation (3), it becomes possible to eliminate the distortion in the X direction.

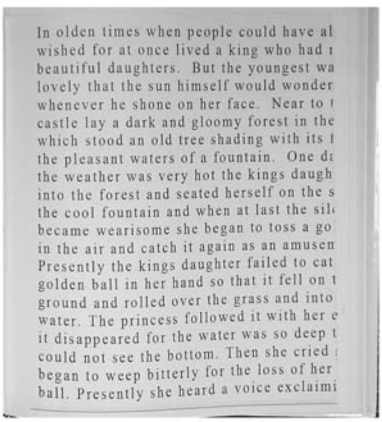
$$x_j = x_i + d_{xji} / s_{ji} = x_i + \frac{X_j - X_i}{Z - f} \cdot f \cdot \frac{X_j^r - X_i^r}{X_j - X_i} = x_i + f \cdot \frac{X_j^r - X_i^r}{Z - f} \tag{3}$$

3.4 Interpolation

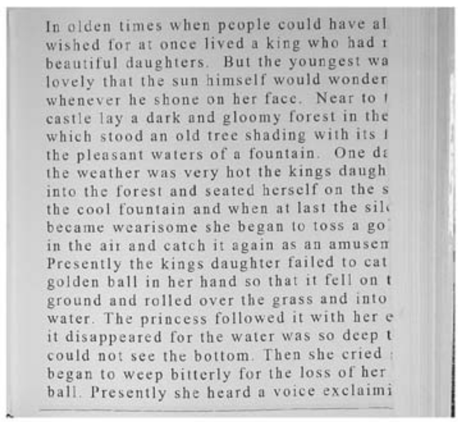
We use linear interpolation here to eliminate the noise introduced through perspective transformation correction. If we let g_i to be the gray level we need to interpolate for the ith point between a gap with length L. Then we apply Equation (4) to yield the results in Fig. 3(d)

$$g_i = \frac{i}{L}(g_{L-1} - g_0) + g_0 \tag{4}$$

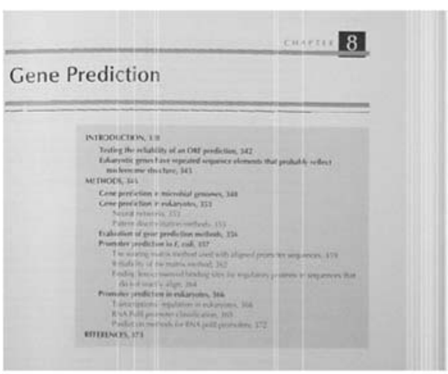
Where g_0 and g_{L-1} are the gray values at the leftmost and rightmost point of the gap.


In olden times when people could have al
wished for at once lived a king who had t
beautiful daughters. But the youngest wa
lovely that the sun himself would wonder
whenever he shone on her face. Near to t
castle lay a dark and gloomy forest in the
which stood an old tree shading with its l
the pleasant waters of a fountain. One da
the weather was very hot the kings daugh
into the forest and seated herself on the s
the cool fountain and when at last the sil
became wearisome she began to toss a go
in the air and catch it again as an amusem
Presently the kings daughter failed to cat
golden ball in her hand so that it fell on t
ground and rolled over the grass and into
water. The princess followed it with her e
it disappeared for the water was so deep t
could not see the bottom. Then she cried
began to weep bitterly for the loss of her
ball. Presently she heard a voice exclaimi

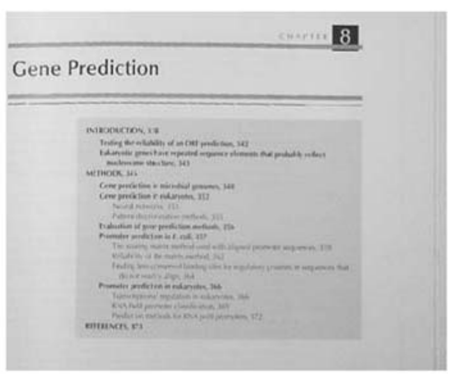
(a) Image after Barrel Correction


In olden times when people could have al
wished for at once lived a king who had t
beautiful daughters. But the youngest wa
lovely that the sun himself would wonder
whenever he shone on her face. Near to t
castle lay a dark and gloomy forest in the
which stood an old tree shading with its l
the pleasant waters of a fountain. One da
the weather was very hot the kings daugh
into the forest and seated herself on the s
the cool fountain and when at last the sil
became wearisome she began to toss a go
in the air and catch it again as an amusem
Presently the kings daughter failed to cat
golden ball in her hand so that it fell on t
ground and rolled over the grass and into
water. The princess followed it with her e
it disappeared for the water was so deep t
could not see the bottom. Then she cried
began to weep bitterly for the loss of her
ball. Presently she heard a voice exclaimi

(b) After Perspective and Interpolation on (a)


CHAPTER 8
Gene Prediction

(c) Image after Perspective Correction


CHAPTER 8
Gene Prediction

(d) Image after Interpolation of (c)

Fig. 3. Barrel Distortion Correction

4 Character Extraction and Recognition

Following image acquisition, the first step involves using the connected components algorithm to find the characters in the image [11]. This algorithm is used to perform a scan of each pixel. Pixels with a value above a set threshold are assigned a label. Any labeled pixels that are 8-neighbors of another labeled pixel are placed in an equivalence class. Using an iterative process, these equivalence classes are combined in order to form each connected component.

Finding Characters within the Image. Logical boundary boxes are drawn around each connected component. The boundary boxes consist of four coordinates that define the boundaries of the object. These coordinates also allow the height, width, horizontal/vertical midpoints, and area of each boundary box to be determined. These measures are used to determine boundary boxes that have been falsely labeled as characters. Boundary boxes with very large or small width to height ratios can be discarded. The discarded boxes correspond to components such as pictures or page separators which do not require classification by the neural network for the time being.

Finding Lines of Text within the Image. Further processing is required in order to group the boundary boxes into lines and then words so that the speech engine will be able to read the text aloud. The first step involves generating a one- dimensional (1-D) histogram of the image from the perspective of the y-axis. The portions of the histogram where the series is at or near zero are used to define the boundaries of each line of text. Boundary boxes with a midpoint above the boundary line are grouped together as a single line.

Finding Words within the Image. In this case, a histogram is generated along the x-axis and pixels between the previously generated boundary lines are counted. Boundary lines are drawn at each zero-crossing, the distance between each pair of lines is determined, and the average spacing between characters is calculated. Boundary line pairs whose spacing is greater than the average spacing correspond to the spacing between words. The boundary boxes are then grouped into words to be passed to the speech synthesis engine after identification by the neural network.

Connecting Disjointed Characters (i and j). Characters such as lower-case ‘i’ and ‘j’ are identified by the connected components algorithm as two separate boundary boxes. One boundary box surrounds the dot portion of the character and another box surrounds the body of the character. Once the boundary boxes have been grouped into lines, each line can be analyzed to determine if any boundary boxes share the same vertical space. When this occurs, the larger boundary box is extended in order to create a single boundary box containing both the dot portion and body of the disjoined character. This process helps the neural network during classification, especially in the case of ‘i’ which is very difficult to distinguish from ‘l’.

5 Neural Network Design

A multilayer feed-forward neural network trained with the back-propagation algorithm is used for character identification. The network consists of three layers of neurons: an input, hidden, and output layer. A size for each layer must be determined before training can begin. Four hundred neurons make up the input layer which corresponds to a twenty by twenty grid that is used to hold each character. This grid size was chosen because it matches the size of an average capital letter of font size 14 in Times New Roman which is considered the default font for the application. The number of hidden units was chosen to be seventy five neurons which allow the network to learn while still maintaining a level of generalization necessary to classify unknown character patterns. The output layer requires as many neurons as there are patterns to classify. There are fifty two lowercase and uppercase letters in the English language. However, when using fifty two neurons for the output layer, the amount of time required for training increased substantially. Instead, each character is represented using a binary coding scheme similar to that used by ASCII code. This change reduced the complexity of the output layer from fifty two output neurons to six output neurons and dramatically decreases the time required to train the network without compromising its ability to identify characters. One final step is required before the neural network can begin training itself for the identification of characters. Each boundary box must be scaled to twenty by twenty pixels so that it can be passed into the input layer of the neural network. After scaling, the network is trained using a training set consisting of each letter in upper and lowercase. The training is stopped once the network is able to successfully identify a cross-validation set consisting of a subset made up of select upper and lowercase letters. The weights obtained through training are saved so that they can later be used to identify unknown characters.

6 Results

The first test case for the trained neural network uses the training set itself as an input. This tests the network's ability to identify a known input pattern. The input pattern for the second test case is a flattened page torn from a book digitized using a scanner. The flattened page contains the same font type and size as the training set. Test case three has as its input an image obtained from a digital camera. The digitization of a page from a book produces a distorted image near the spine due to the page's curvature. The results from this test case reveal the need for additional image processing to correct the curvature distortion. Test case four applies the perspective transformation correction. This correction improves the accuracy of identification only slightly. Further distortion is being caused by the barrel of the digital camera. The barrel distortion causes increasing image curvature as the distance of the pixels from the center of the image increases. The results of classification once the barrel distortion correction is applied are shown in Test Case five. Preliminary results indicate that the order of distortion corrections does not affect the accuracy of the outcome (whether barrel distortion correction follows perspective transform correction or vice versa). The final test case integrates the spell check function of Microsoft Word to correct classification errors, see Table 1.

Table 1. Results of the different test cases

Test Cases	Correct Classifications	Incorrect Classifications	Accuracy
Case 1: *Training Set*	53	0	100%
Case 2: *Scanned Torn Page*	725	1	99.98%
Case 3: *Uncorrected from Book*	353	372	48.68%
Case 4: *Perspective Correction*	400	325	55.17%
Case 5: *Case 4+ Barrel Correction*	650	75	89.65%
Case 6: *Case 5+MSW Correction*	673	52	92.82%

The processing time for reading a page of the book is provided below. Preliminary results prove to be encouraging in contrast to an average reading time of sighted readers. The algorithms we have developed in terms of dealing with distortions and in terms of the neural network training will have to be refined for better accuracy and remodeled to enhance the processing time for a faster reading outcome. We need to emphasize that at this stage of the research, images and tables within the pages processed are discarded within the context of the proposed algorithm. In the comparison between human subject read times versus the automatic document reader's the total time is from digitization to vocalization as shown in Tables 2 and 3.

Table 2. Time consumed in seconds through image pre-processing

Image Size	Barrel correction	Curve Detection	Perspective correction
1944*2592	11	10	18

Extension Correction	Interpolation	Total
18	8	65

Table 3. Overall Read Time Results

Human vs. Document Reader	Time (seconds)
Human Subject (Averaging 4 Subjects)	56.25
Automated Document Reader	23 (ANN) + 65 (Image Preprocessing)

7 Conclusion

This research introduced a new book reader system that is inexpensive and effective in terms of reading accuracy. The main contributions are in the mathematical developments that address the critical issues of page curvature, perspective transformation, and barrel distortion due to image capture; and in the development of a neural network algorithm that address the problem of character recognition with the abilities for online retraining based on prior errors made in the first pages that were read, and in embedding the spell check function of Microsoft Word to enhance the accuracy of the final read-out. Classification accuracy of this book reader system is further enhanced through the integration of two additional modules. The first module begins by identifying each of the unique characters within a page. Each unique character is then passed to the previously trained neural network for identification. Incorrect classifications initiate the network to retrain until it can successfully classify the previously mislabeled characters. This learning process continues iteratively throughout subsequent pages until the network is able to correctly classify nearly all of the characters within the book. The second module uses the Microsoft Office spell check function to determine the correct spelling of each misspelled word. The first spelling chosen by the spell check function is assumed to be correct.

Acknowledgements

The authors gratefully acknowledge the support from the National Science Foundation grants: EIA-9906600, CNS 0426125, NSF IIS-0308155**,** and HRD-0317692.

References

1. Amin A., Fischer S.: A document skew detection method using the Hough Transform, Pattern Analysis and Applications: 3 (3), pp. 243-253, Sept. 2000
2. Ishitani Y.: Logical Structure Analysis of Document Images Based on Emergent Computation:, IEICE Trans. on Information and Systems (8), pp. 1831-1842, August 2005
3. Nagy G.: Twenty Years of Document Image Analysis in PAMI: IEEE Transactions on Pattern Analysis and Machine Intelligence 22 (1), pp. 38-62, Jan. 2000
4. Muller K.R., Mika S., Ratsch G., Tsuda K., Scholkopf B.:An Introduction to Kernel-based Learning Algorithms, IEEE Trans. on Neural Networks 12 (2), pp. 181-201, March 2001
5. Rao S.S.: Electronic books: their integration into library and information centers:, Electronic Library 23 (1): 116-140 2005
6. Alonso M., Barreto A., Cremades J.G., Jacko J.A., ad Adjouadi M.: Behaviour & Information Technology 24 (3): 161-173 May-Jun 2005
7. Adjouadi M.: "A Man-Machine Vision Interface for Sensing the Environment", Journal of Rehabilitation Research and Development, Vol. 29, No. 2, pp. 57-76, 1992
8. Adjouadi M., Riley J., Candocia F., Andrian J., Sumargo H.: An Augmented Computer Vision Approach for Enhanced Image Understanding: Journal of Rehabilitation Research and Development, Vol. 32, No. 3, pp. 264-279, October 1995
9. Pajdla T., Werner T., Hlavac V.: Correcting Radial Lens Distortion without knowledge of 3-D Structure: Technical report TR97-138, FEL CVUT, Prague, Czech Republic, 1997
10. Vass G., Perlaki T.: Applying and Removing Lens Distortion in Post Production: Colorfront Ltd., Budapest, 2003
11. Casey R.G., Lecolinet E.: A Survey of Methods and Strategies in Character Segmentation:, IEEE Trans. on Pattern Analysis and Machine Intelligence 18 (7), pp. 690-706, Jul. 1996

RoboBraille – Automated Braille Translation by Means of an E-Mail Robot

Lars Ballieu Christensen

Sensus ApS, Godthåbsvej 18, DK-3400 Hillerød, Denmark
lbc@sensus.dk
www.sensus.dk

Abstract. As society becomes increasingly dependent on literacy, the problems of textual information inaccessible to print-impaired people are likely to grow. This paper discusses the problems of decentralised, user-driven Braille translation and proposes an alternative: The centralised, email-based RoboBraille service capable of translating to and from contracted Braille, including any pre- or post processing steps required to convert between document types, formats and character sets. As such, the RoboBraille service attempts to solve a universal problem as it makes textual information accessible to people who would otherwise find it inaccessible due to disability or reading difficulties. Originally a Danish service, a pan-European consortium is currently validating RoboBraille in six European countries with financial support from the European Commission.

1 Introduction

In Denmark, software to translate to and from contracted Braille in multiple languages has been available since the mid 1980s [1][7]. Although the systems are fairly easy to use, fast, accurate, well-promoted by the support system and available free of charge as downloads from websites, they are not being widely used amongst teachers, Braille readers and others with a need to produce contracted Braille. Why? Traditionally, translating documents into contracted Braille is a time consuming process that requires a wide range of different skills: In addition to mastering the Braille translation software, translators must understand how to handle different document types, character sets and formats. Once translated, the translator must also know details about the Braille device on which the text is to be displayed or rendered. And since Braille translation is a niche with limited resources, software is constantly being updated with software patches. For professional Braille translators, these issues may not pose a problem. However, for the occasional translator – e.g., a primary-school teacher with an integrated blind pupil or a blind Braille reader – they do.

This paper discusses an alternative to the decentralised, user-centric Braille translation systems used widely to produce Braille at varying levels of contraction. Based on experience from developing automated Braille translation solutions during past 20 years, the RoboBraille service is a centralised, e-mail based translation agent that automates the translation process, including any pre- or post processing steps required to convert between document types, formats and character sets. Since

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 1102 – 1109, 2006.

RoboBraille is based on e-mail, the solution is platform independent and the only skill needed to use the service is the ability to send and receive an email with a document attachment. The RoboBraille architecture is based on standard internet technologies and can be managed centrally. Consequently, the solution is highly scalable, always up to date and can be operated by a minimum of efforts. In addition to translating to and from contracted Braille, RoboBraille is capable of producing audio (mp3) files from submitted documents using integrated text-to-speech synthesisers and audio compression software. The service is being offered free of charge to all non-commercial users.

The RoboBraille service attempts to solve a universal problem: It makes textual information accessible to people who would otherwise find it inaccessible due to disability or reading difficulties. As society becomes increasingly dependent on literacy [3], the problem of textual information inaccessible to print-impaired people is likely to grow. From a conceptual point of view, the RoboBraille is a universal solution to a universal problem. By adapting the service to include the specific Braille translation and text-to-speech capabilities of any given language, the service will be of interest to governments, commercial organisations, interest groups and print impaired people alike irrespective of region or nationality. Using standardised internet technologies to interact with the service supports the notion of a fully pan-European or global service. Originally a Danish service developed by the author in close collaboration with Synscenter Refsnæs, the national Danish Centre for Visual Impairment, Children and Youth (see www.synref.dk), the European Commission has recognised the potential of RoboBraille, and has retained a project aimed at validating the service in five other countries in 2006 and 2007 as part of the eTEN programme. Headed by Synscenter Refsnæs, the eTEN RoboBraille consortium includes seven European organisations.

2 Challenges of De-centralised Braille Translation

Although electronic versions of texts are becoming widely available (or can be scanned and converted to text using OCR software), many documents are kept in older formats or character sets such as the old MS-DOS code pages. Similarly, although certain document formats such as Microsoft Word, Rich Text Format (RTF) or Portable Document Format (PDF) may be suitable for producing visual, printed copies, they may not be supported by the Braille translation software. Finally, since few (or too many) standards exist for how to layout the Braille character set on Braille devices and embossers, few Braille devices and embossers appear to share a character set. Consequently, traditional decentralised Braille translation requires the translator to obtain an electronic copy of a particular text, to convert it into a format and a character set suitable for the Braille translation software, to translate the text into contracted Braille and to convert the resulting text in accordance with the character set of a particular Braille device or embosser. During the process, it may even be necessary to chop up the text into smaller parts in order to fit it into the memory of a Braille device.

Needless to say, the person who only occasionally needs to translate a document into (or from) contracted Braille has a hard time. And should he or she succeed, chances are that the Braille translation software is out of date and the resulting Braille text flawed. Although we have not conducted a formal survey amongst occasional Braille translators in Denmark, our experience with these users – dating back to the mid 1980s – has confirmed our thesis: Decentralised, user-centric Braille translation is too cumbersome for the occasional user. As a result, teachers of integrated blind-born children in the comprehensive Danish *Folkeskole* abandon contracted Braille in favour of uncontracted Braille or even synthetic speech. Likewise, the blind themselves remain dependent on centralised Braille production sites, who can muster all the required skills, in order to gain access to Braille renditions of textual material.

Far worse than the lack of independence is the erosion of basic Braille skills amongst the blind that this situation causes [2]. Over the course of the past 30 years, Braille literacy has shown a dramatic decline. In America [4] and the UK [5], alarming statistics have been published and although less significant, similar trends are reported from other countries (e.g., [6]). Many fail to recognise the importance of Braille: It is difficult to learn as a visually impaired person, difficult to read and understand as a sighted person, it is costly and time-consuming to produce and Braille devices such as Braille printers and Braille displays are expensive. In many cases, speech synthesis appears to be an attractive alternative. However, a symbolic written medium is as important for the blind as it is for the sighted. Braille is a fundamental means of communicating and plays a significant role in the process of intellectual development: It is so much more effective to be an active reader than a passive listener. Furthermore, Braille literacy is an integral part of the personal identity for visually impaired people [5]. In general, lack of Braille skills amongst the blind equals illiteracy – a rather serious issue in the information society.

3 The Process of Resolving the Issues of Complexity

Although the ultimate solution to the issues of complexity in traditional, decentralised and user-centric Braille may be simple – to the degree of seeming obvious – the process of reaching the solution has been long and circuitous. The Danish Braille translation solutions have evolved alongside the general technological development and the adaptation of information technology amongst the visually impaired over the course of the past 20 years. At each stage in this development, barriers have been overcome, only to be replaced with new obstacles. In some case, solutions have been developed to address particular problems. In other cases, solutions have materialised as a result of the general technological development and adaptation.

In the mid-1980s, the main issues facing Braille translation was the availability of translation software on affordable computer systems combined with the lack of electronic versions of textual material. As PCs and scanners became widely available and commercial text editing became electronic, the issues changed and became a matter of distribution, Braille consumption and intellectual property rights. Around 1990, eBooks for the blind became a reality with the establishment of diskette-based distribution and electronic bulletin boards (BBS); these services were eventually replaced by email and the web. Likewise, Braille readers were offered Braille displays or personal Braille printers as alternatives to paper-based Braille. During the past decade, Braille translation software was introduced to the Braille readers and those preparing ad-hoc material, electronic web-based libraries for blind were established and the Braille-enabled computers improved [7].

Alongside the technology adaptation amongst the visually impaired, information technology was adopted by the general public. From being a rarity 10 years ago, most households today have one or more computers, and the computers are connected to the internet – in many cases through broadband connections. Most people are familiar with emails and electronic documents, and make use of the web for a variety of purposes [8].

20 years ago it would not have been possible. In the current technological climate, however, one solution to the issues of complexity in traditional, decentralised and user-centric Braille translation was obvious: RoboBraille, an email-based translation agent.

4 The Centralised Alternative

Although RoboBraille does not claim to resolve all the challenges of Braille, the service does ease the process of producing high-quality contracted Braille. Braille remains difficult to learn as a visually impaired person, it is difficult to read and write as a sighted person, and Braille devices are costly. However, the actual process of transforming textual documents such as books, articles, bank statements and medical instructions into Braille is significantly simplified. As a positive side-effect, the RoboBraille service can also be used to produce other forms of Braille – e.g., visual Braille for pharmaceutical companies with a need to emboss tactile Braille on pharmaceutical packages and labels – as well as audio files with synthetic speech.

The RoboBraille service is an email-based translation service capable of translating documents to and from contracted Braille and to synthetic speech. Users submit documents (e.g., text files, Word documents, HTML pages) as email attachments. The translated results are returned to the user via email. The user interacts with the RoboBraille service by sending emails to specific email accounts. The main email accounts are listed in Table 1 below:

Table 1. Main RoboBraille translation processes

Email account	Process
`ottepunkt@punktskrift.dk`	Translate document into contracted eight-dot Braille.
`sekspunkt@punktskrift.dk`	Translate document into contracted six-dot Braille.
`fuldtekst@punktskrift.dk`	Translate (presumed contracted) document into standard text.
`filepart@punktskrift.dk`	Partition document into smaller parts.
`converter@punktskrift.dk`	Convert from one document format to another.
`eksport@punktskrift.dk`	Convert to character set of a particular Braille device.
`tale@punktskrift.dk`	Translate document into synthetic speech.
`sb4admin@punktskrift.dk`	Administrative account for status and updates.

The service has been running as a prototype in Denmark since August 2004; in December 2005, the number of processed requests exceeded 12,000.

The current version of the RoboBraille service supports a range of popular document formats, including standard text, HTML, Word and Rich Text Format (RTF). Prior to translation, Word and RTF files are converted into text. Depending on the size of the file, the traffic and server workload, a result is typically returned to the user in a matter of minutes of submitting a request for translation. Figure 1 below illustrates how a user interacts with the system in order to have a document translated into contracted Braille:

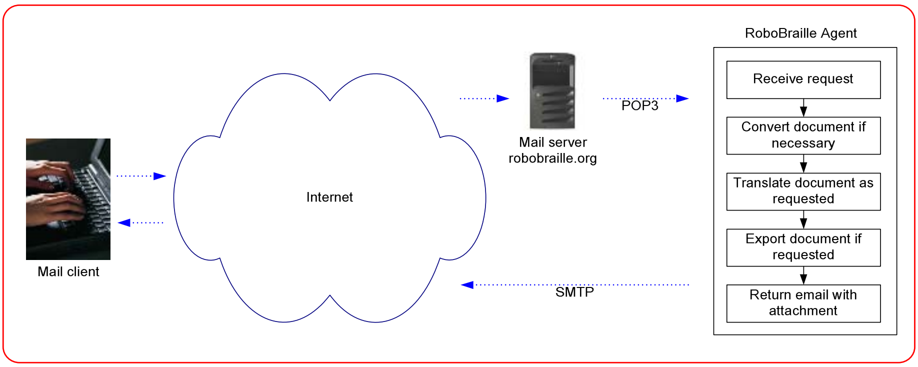


Fig. 1. The RoboBraille Braille translation flow

The translation system assumes that the source document is written in the standard Windows character set for Western Europe (ISO 8859/1). Furthermore, the system supports automatic conversion of older ASCII documents to Windows text files. Once translated, the document is returned in OctoBraille, a Braille adaptation of the standard ISO 8859/1 character. Since few Braille devices share the same character set, the translation system can convert the translated document into a range of different formats to accommodate Braille note takers and embossers.

Likewise, the user may request a document be translated into synthetic speech. The process is similar to that of Braille translation, although some of the steps are different. First, the translation system translates an attached document into a WAVE file. WAVE files are rather large and unsuitable for transmission via the Internet. Therefore, the WAVE file is subsequently encoded and compressed into an MP3 file. The resulting audio file is copied with a unique name to the web server using FTP, and a link to the file is returned to the user.

In order to provide a scalable and high-performance solution, the RoboBraille service is based on a two-tier architecture consisting of a server layer and an agent layer, respectively (see figure 2 below). The server layer includes a mail server for receiving incoming requests and a web server for delivering audio contents. The agent layer consists of a range of identically configured desktop computers running the RoboBraille software package on Microsoft Windows.

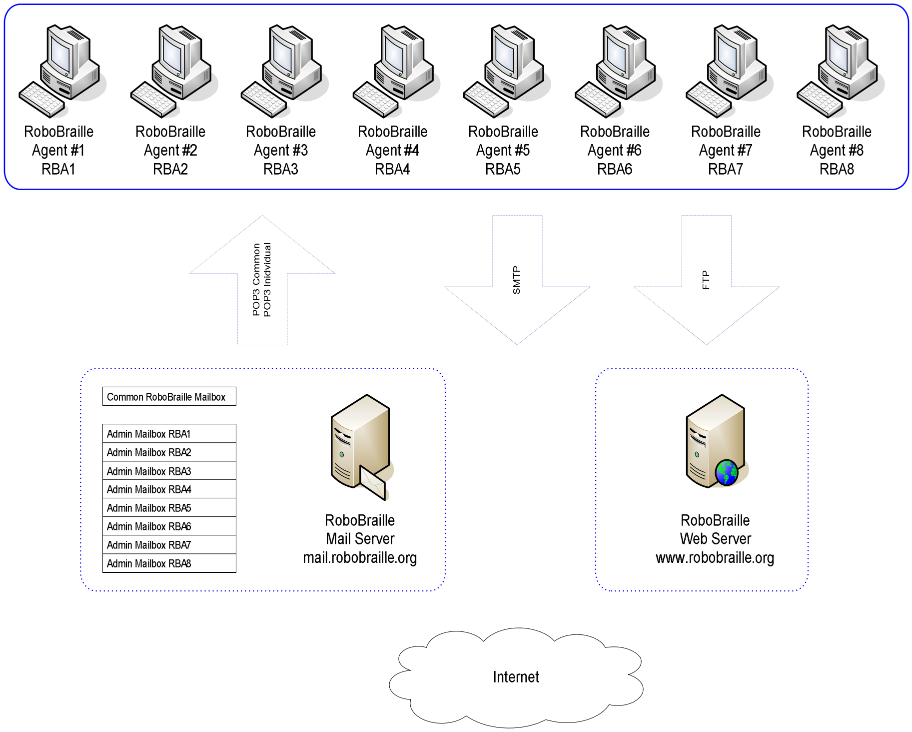


Fig. 2. The Overall RoboBraille Systems Architecture

Incoming requests are received by the mail server and subsequently divided into two different categories:

1. Translation requests: These include all requests for translation of attached documents into either Braille or synthetic speech. All translation requests are placed in the common RoboBraille mailbox, irrespectively of which translation service is requested. All RoboBraille agents read from this mailbox using the POP3 protocol as part of a scheduled poll sequence. The poll interval is typically set to 1-2 minutes. With a poll interval of 1 minute and 8 RoboBraille agents, the RoboBraille service will have a theoretical capacity of servicing 60 * 8 = 480 requests per hour. The common RoboBraille mailbox serves as a first-in-first-out queue. Once a translation request has been read from the common RoboBraille mailbox by a RoboBraille agent, it is deleted from the mailbox. Further processing of the translation request is managed by the individual RoboBraille agent.
2. Administrative requests: These include a number of different request types sent to the RoboBraille service: Request for status, update of data tables, and update of software. As such, these are the requests used to manage the RoboBraille service and ensure that all RoboBraille agents are up-to-date. Rather than delivering these requests to the common RoboBraille mailbox, administrative requests are copies to an individual administrative mailbox for each RoboBraille

agent. All RoboBraille agents read from its administrative mailbox using the POP3 protocol as part of a scheduled poll sequence. Once an administrative request has been read from the individual administrative mailbox, it is deleted from that mailbox. Only selected users can issue administrative requests. These administrative requests allow administrators to check the status and to update the data tables and software modules of all RoboBraille agents simply by sending emails to the RoboBraille service. All RoboBraille agents automatically reboot every night at midnight, loading all updates that may have been mailed as administrative requests during the past 24-hour period.

The RoboBraille software package includes the RoboBraille Automated Mail Responder (AMR) module including software components for graphical manipulation and audio compression, Sensus Braille 4 Braille translation engine, text-to-Speech synthesizers and Microsoft Office Automation components to handle Word and RTF documents.

As the communication between the mail server, web server and the RoboBraille agents are based on standard Internet communication protocols (POP3, SMTP, HTTP and FTP), the individual computers do not need to be located on the same physical location. However, in order to make efficient use of systems administration resources, limit the amount of resources required to keep the systems updated and reduce the risk of unscheduled down-time, all machines are maintained centrally at two separate locations in Denmark.

Each RoboBraille-enabled desktop computer is configured to serve several languages (Braille and text-to-speech translation). Likewise, multiple RoboBraille-enabled desktop computers will be available to serve the same language. As each RoboBraille-enabled desktop computer is running independently of the others, the architecture is truly scalable, and the ability to add new agents will be used to ensure the service level and balance the workload. Experience from the Danish RoboBraille service suggest that deploying identical RoboBraille agents all capable of processing translation requests is a simple means to ensure efficient load balancing and availability.

5 Europe and Beyond

As a universal solution to the problem of inaccessibility to textual information amongst print impaired people, the potential of the RoboBraille solution is substantial. Already, a pan-European market validation project has been retained for funding by the European Commission under the eTEN programme. Once validated in Ireland, the United Kingdom, Italy, Portugal and Cyprus, the RoboBraille consortium expects to maintain the service in the six countries and deploy the service further into other interested countries. Through dialogues with relevant government agencies, NGOs and institutions for the blind, partially sighted, dyslexic and other print-impaired people, and by maintaining a high level of visibility through participation at conferences and press coverage throughout Europe, the RoboBraille consortium will invite other countries to join the service. The aim is to add between two and five new countries/languages per year until all interested EU Member states have joined the service. Several countries outside the EU are also expected to join the service.

Critical to the further expansion of the RoboBraille service is the adoption amongst public and private commercial organization. For this purpose, the RoboBraille service will be developed into a step-stone technology, allowing organisations to communicate electronically with citizens, clients and customers in formats alternative to plain text. The term step stone is used to describe the scenario where a piece of information – e.g., a tax return form, a bank statement, a letter – is mailed to someone via the RoboBraille service. Using the RoboBraille service as an intermediary, the originator of the information may be able to send textual information in a multitude of alternate formats such as Braille, speech and enlarged print. Furthermore, the ability to produce visual Braille for companies who need to emboss tactile Braille on packages, lables, bottles, etc., will be enhanced.

References

1. Christensen, L.B.: Multilingual Two-Way Braille Translation. From Klaus, J, et al: Interdiciplinary Aspects on Computers Helping People with Special Needs, Österreichische Computer Gesellschaft/R. Oldenbourg Wien München (1996)
2. Christensen, L.B.: The Importance of Information Technology for Visually Impaired Children and Youngsters and the Expectations for Future Development. Conference proceedings. ICEVI Europe (2000)
3. Literacy in the Information Age. Final Report of the International Adult Literacy Survey. OECD (2000)
4. Schroeder, F.K.: Perceptions of Braille Usage by Legally Blind Adults, Journal of Visual Impairment & Blindness, May-June (1996)
5. Bruce, I., McKennell, A., and Walker, E.: Blind and partially sighted adults in Britain: the RNIB survey, RNIB/HMSO (1991)
6. Kahlisch, T., and Lötzsch, J.: Services for the Blind and Partially Sighted in Germany, Technischen Universität Dresden. From Congress on Rehabilitation of the Disabled, Dubai, United Arab Emirates, Oct. 6.-9. 1996 (1996)
7. Christensen, L.B.: Applying Information Technology as an Intelligent Interface for the Blind, University of Roskilde (1996)
8. Ministry of Science, Technology and Innovation: Information Society Denmark ICT Status (2005)

Accessing Music Notation Through Touch and Speech

Ben P. Challis

Department of Arts and Media,
University of Glamorgan, Pontypridd, Wales, UK CF37 1DL
bchallis@glam.ac.uk

Abstract. A system for the non-visual presentation of music notation is described. Derived from an earlier test-platform for the establishment of a set of principles for tactile communication, the Weasel notation system has been redesigned to facilitate both music-reading and composition tasks. Using tactile overlays on a touchpad, speech output/input and audio-playback, the system allows the reader to browse and interact with music notation in a way that is appropriate to the majority of common music reading and writing tasks.

1 Introduction

The standard notation for western music is Common Music Notation (CMN). Although other notations exist[1], and are in relatively common use, it is CMN that is accepted as standard. Indeed, on opening a typical school or college music-textbook, it is almost certain that CMN will be the notation system in use for describing musical excerpts. Whether CMN is the most efficient means of communicating music-based information is debatable; after all, this is a notation system that has evolved across a number of centuries adapting along the way to a general increase in complexity of both orchestration and instrumental performance. Putting any potential limitations to one side, CMN is almost universally accepted as the standard form of western music notation, not least because it can be applied to nearly all musical instruments. However, the strength of CMN can be found in its visual richness and it is this same 'strength' that effectively excludes blind musicians from accessing it easily.

2 Visual Properties of CMN

Music notation is graphic, lending itself to the notion of 'filtering' out that which is not necessary to a specific task. For example, at a high level, an instrumentalist who is new to a particular musical-arrangement is able to breakdown the 'whole' into logical

[1] Most fretted stringed-instruments (e.g. guitar, mandolin, violin) can employ a system of tablature, where a diagram of the fretboard is used to show finger positions on individual strings. Some wind instruments (e.g. flute, whistle) can use a similar style of finger-chart system to show a row of pitches represented by 'closed' or 'open' holes on small diagrams of the instrument's body.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 1110 – 1117, 2006.

sections. Various graphic components of CMN (e.g. double barlines, repeat signs and sometimes rehearsal marks) can assist with identifying logical sections to work with. At a much lower level, CMN allows the identification of pitch and duration of individual notes but can also include other instructions, some symbolic and some textual, to describe how to play the notes (e.g. loud, soft, fast, slow, detached etc.). One particularly strength of CMN is that a single piece of notated music can be used by musicians of differing abilities for a range of purposes. Where one individual may wish to learn an entire piece note-by-note, another individual may be familiar enough with the rhythm of the melodic line that only the step-by-step change in pitch is of interest; another may only be interested in the underlying chord progression. Typically, all of these tasks, and many others, can be achieved from the same piece of music notation.

3 Non-visual Approaches to Music Notation

Although there are a number of non-visual alternatives to CMN, these do have limitations when compared to the different levels of presentation of data afforded through graphic music notation. Typically these systemss fall into one of two categories: tactile presentation or speech access.

3.1 Tactile Systems

Tactile systems include Braille Music, Moon and tactile CMN (enlarged). Braille music is without the doubt the most common tactile approach to music notation in use with many libraries and services offering access to large quantities of music in Braille format. However, Braille music does have its problems, not least that the number of blind people who actually use Braille is far lower than might be expected [1]. In addition, Braille music is not as standardised as would be hoped, uses alternate characters for music letters to those used in conventional spelling and needs to be read in series.

Moon music is less complete as a system and enlarged CMN has the obvious limitation of how much information can be displayed within an easily accessible area. Purely tactile systems are also more convenient for reading-tasks than compositional-tasks. Significantly, none of these systems offer a convenient structural overview; the focus is mainly at event-level. Neither can they offer any level of filtration i.e. a reader needs to read everything before deciding what information is of direct use within a given task (e.g. learning fingering for a keyboard part).

3.2 Spoken Systems

The most basic system for displaying music notation through speech is the 'Talking Scores' system as used by the RNIB[2]. A piece of music is described on a bar-by-bar, event-by-event basis using recorded vocal descriptions. As with tactile approaches,

---

[2] Royal National Institute for the Blind, UK, (`www.rnib.org.uk`).

this generally involves the reader listening to the whole recording or at least listening to more detail than might be required. This approach has been enhanced considerably through the introduction of a computer-based approach to implementing the Talking Scores protocol [2].

Perhaps one of the most popular notation packages for CMN is Sibelius[3]. Script-based access to this system has recently been made available[4] and similar scripts are now also available for the Lime music notation package. These systems effectively provide access to spoken descriptions of menus and graphic symbols for use with a screen-reader (e.g. JAWS or HAL).

Again, these systems do not provide a facility for a structural overview, and there are compromises in usability depending on whether the task is one of reading or composing.

3.3 Playing-by-ear

It is important to clarify the difference between *playing-by-ear* and *learning-by-rote* as both terms are used quite freely in music education. *Playing-by-ear* involves the learner listening to and then copying a piece of music, section-by-section. *Learning-by-rote* involves one individual demonstrating to another how to play a particular piece. Here, the learner is 'shown' how to play each section although it may be that an element of *playing-by-ear* is involved.

Whereas *playing-by-rote* is unlikely to be appropriate for many blind music-learners, learning to play by-ear can have a useful place. However, the learner's sense of pitch and timing needs to be particularly accurate; this is often not the case with learners in the early stage of music education.

4 The Weasel Project

The Weasel[5] project is a method of providing computer-aided access to a system of music notation that is part-tactile and part-electronic in presentation. In essence, a tactile-overlay is used in conjunction with a touch-pad to provide an interactive overview of a piece, or page, of music. The overlay is used to display the significant structural elements within the piece and the rest of the musical information is displayed using a combination of speech synthesis and audio-playback.

4.1 Tactile Overlays

It is important to appreciate that the overlay itself contains absolutely no specific note information. If a user were to explore the tactile component alone, very little could be identified about the piece being displayed beyond the number of bars to a line (and lines to each page), whether there are identifiable sections (or similar rehearsal points) and

---

[3] Sibelius Software Ltd, UK, (`www.sibelius.com`).

[4] Sibelius Speaking from Dancing Dots, PA, USA, (`www.dancingdots.com`).

[5] Weasel was first conceived during a research project into establishing principles for tactile communication at the University of York, UK.

whether any bars or sections are repeated. Essentially, the tactile overview is a navigational aid. Importantly, terms of reference such as 'the-bottom-line' or 'bar 2 line 3' or 'from the repeat mark at the end of line 1' all remain valid in a way that is not immediately apparent in other systems. This is an important aspect as music performance and learning often involves other musicians or music-learners. It would be undesirable to create a system that is efficient at communicating to the individual but that does not facilitate a common understanding of 'location' within the piece; this would not be truly inclusive.

By navigating around the tactile overview the user can interrogate the music in various ways that effectively allow for individual levels of requirement (similar to those described earlier with reference to CMN). The information required can then be accessed in further detail using either speech or audio-playback. This combination not only allows for information to be filtered such that the reader can focus on differing levels of complexity but also allows the user to employ a preferred method of learning – by-ear, by spoken-description, and potentially by Braille (if a refreshable display were to be employed).

In its original versions [3], [4], Weasel used vacuum-formed PVC overlays such that differing levels of height could be achieved. Whilst this was appropriate to the original investigation, the production of such overlays is time-consuming and impractical in terms of production by an average user; the equipment required is expensive and individual moulds are required for each overlay.

To achieve a more practical system, Weasel has been completely redesigned such that in addition to tactile overlays and speech-output, speech input is being incorporated. This can effectively free the user from the computer-touchpad, whilst still allowing the system to be operated. In terms of music, this has obvious practical advantages for instrumentalists who will normally use both hands to play their instrument.

Overlays are now prepared using 'swell' paper with a single overlay representing a single page of music. Using raised lines and symbols, very basic structural information can be presented to the user. Abandoning the use of PVC overlays has meant that the use of 'change-of-height' as a method of encoding instructions has become difficult to achieve. Swell paper can only practically provide a single change in level with a maximum height of around one millimetre. Although it was demonstrated in earlier experiments with tactile-overlays that differing changes in height could be useful it was decided that the additional richness achieved was outweighed by the complexities involved in production. An aim of the current project has been to make the platform available for use with the minimum of additional effort.

4.2 Interaction

On the overlays, guidelines are employed to trace across bar areas, with additional symbols being used to represent other structural elements that are of use (e.g. repeat marks, rehearsal points). Interacting with the overlay by pressing on a 'bar' area will force the system to describe the contents of the bar using either speech, audio-playback or a combination of both if appropriate. The information described depends upon the current preferences of the user (e.g. pitch, duration, dynamics, fingering, chords etc.). Areas of bars (sections) can be selected for playback such that rehearsal becomes more manageable.

The majority of actions can be made directly. For example, pressing the barline symbol at the beginning of the piece will inform the user of a number of essential items of information (tempo, key, time-signature, starting dynamic etc.), these can be accessed one-at-a-time by pressing the symbol repeatedly. Pressing the symbol for any other barline will 'speak' the bar-number for that current bar. Pressing repeatedly will progress through any other important pieces of information that might be present at that same location, a change in dynamics for example. Pressing the symbol for a repeat mark will inform the user of where to return to within the piece to begin repetition.

Currently, the system relies on an Intellikeys[6] touchpad as these are readily available and are typical of the technology in use in special-needs environments. The surface resolution of this device is limited but can be regarded as being similar to the size of a fingertip. With this in mind, it would be difficult to 'draw' a smooth line using one of these touchpads, however, the interactive tasks used within the system are generally simple 'single-click' operations. The overlays are being designed using established principles for the design of interactive tactile diagrams [5], [6].

4.3 Controls

Besides directly interacting with tactile components within the overlay, the user can also navigate through a simple menu-system to access various features important to tailoring the display to suit individual needs. As described earlier, it is often the case that a music reader will have one or more specific tasks in mind. Rather than reading everything that is present within the score, it might be that only pitch or duration information is of immediate use. Alternatively, it might be that instrument specific nuances are of interest e.g. bowing or fingering. In addition, besides filtering out that which is irrelevant to a given task, the user will also need to make decisions on how best display that information if more than one possibility exists. For example, pitch information can be spoken but can also be heard as melodic notes. Where one user may prefer to read by-ear, another may prefer to be informed of the letter name (and accidental if present). Equally, rhythms can be described using speech or played in context (perhaps with pitches) or extracted as pure rhythm (pitches removed and substituted with a 'click' sound). Regardless of preference, the user needs to be able to quickly alter their preferences as they move from task to task. With this in mind, the controls for making such changes are embedded within the main display. This avoids the user needing to move between the overlay/touchpad and the computer keyboard unnecessarily.

The controls employ a simple design that has been shown to work effectively in previous versions of Weasel. In essence, three fingers from each hand can be located in two sets of three 'rings' at the bottom of the overlay. The left set of controls navigates through the menu headings and the right set navigates within the list belonging to the current selected menu. Pressing the centre 'ring' of either set confirms what the current

[6] Intellitools, CA, USA, (`www.intellitools.com`).

selected menu or item is. The output is always speech although audio-reinforcement is used to help distinguish between the 'menu' and 'item' controls.

4.3 Output

For music-audio output, Weasel sends standard MIDI message to be received by a typical general MIDI sound-module. This reduces the demand for real-time sound-processing that can impact upon smooth speech output (and input). Many computer soundcards are already equipped with basic MIDI sound modules. For speech synthesis, the system takes advantage of the TTS (Text-To-Speech) and SR (Speech Recognition) engines that now form part of the Windows XP operating system. In all, the system avoids using specialist hardware beyond that which is typically already available in most special-needs education-centres.

4.4 Input

The aim of the project has been to explore approaches for displaying music notation in a non-visual format. With this in mind it has been essential to have an efficient method of inputting music data into the system. Although the system is to be expanded in the future to allow composition, not every music-reader will only wish to read their own compositions. Therefore, it is important for there to be easy access to a library of scores. There are limited solutions to this.

Using standard MIDI files would present a large library of music to most users as this format has been in popular use since the 1980s. However, the protocol for the use of these files has never been one of providing a substitute to music notation. Where MIDI might excel in terms of storing pitch, duration and velocity information it provides almost no method of storing many other musical concepts (fingering, position, performance symbols etc.). MIDI files are used as input mechanisms for some software packages (e.g. Sibelius and Finale), however, the function of these packages is perhaps closer to that of typesetting.

In contrast, ABC is a text-based approach to music notation that has been in use since the early 1990s. Originally designed for the sharing of folk music on the internet, ABC has rapidly become widespread in its use and offers a more complete approach to inputting and storing music notation than that offered by MIDI. As all the music information and performance instructions are available this is a good medium for both the storage and presentation of musical information. Indeed, the Dutch Federation of Libraries for the Blind (FNB) has integrated this file protocol as the basis for its *Spoken Music* project; a computer-aided equivalent to the Royal National Institute for the Blind's *Talking Scores* (RNIB, UK). It would perhaps be premature to suggest that the ABC protocol is, or ever will be, the standard text-based music notation system but it is currently the most common and is well documented. With these aspects in mind, coupled with its usage within the Spoken Music project, ABC has been adopted as the intermediate file system within the Weasel project.

5 Future Work

The Weasel project has aimed to provide a flexible system for the reading and composition of music-notation. However, the majority of the focus has been on the creation of a working ‘engine’ for presenting spoken-descriptions and audio-playback of extracts of music. There are separate issues that arise when considering the nature of compositional tasks in a non-visual format. In addition, it has also become apparent that the project may be of considerable use to some users who are not visually-impaired.

5.1 Music Composition

In its current state, Weasel responds to input from a computer keyboard, an Intellikeys touchpad and speech. It also outputs both speech and audio-playback. However, a number of overlay-designs are under consideration and are yet to be tested thoroughly to discover which is most effective as a navigational aid. A further aim within the project is to adapt Weasel such that it facilitates basic composition but it remains to be seen whether a system can be developed that will allow an overlay to be constructed part-by-part. Ideally, a fully interactive dynamic tactile display would fit this purpose but the design of such a device is beyond the scope of this project. It is also unlikely that such a device will be available for some time yet, certainly at an affordable level. However, it is conceivable that for short-term test purposes a system of blank template-overlays could be used. These will use a set number of bars-to-a-line and lines-to-a-page. The user will be able to compose tunes within this framework, accessing bars for recording and editing note or performance information. When the process of composition is completed, the layout can be modified such that the structure becomes more apparent and logical to the reader.

5.2 Music and Dyslexia

Although Weasel has been developed with visually-impaired users in mind it now appears that there may be another user-group who could benefit. Personal experience as a music lecturer has shown that students with dyslexia will generally also experience difficulty in the general comprehension of written-music. These students can often have good listening skills and may prefer to learn music by-ear rather than by reading notation. However, learning to play music by-ear does not lend itself to composing and arranging parts for other musicians; this can make academic progress appear to be slower than might be expected. It is conceivable that this same system may provide an approach to reading and writing music that is perhaps more accessible than more traditional approaches.

6 Conclusion

The Weasel project has attempted to provide a cost-effective solution to a number of accessibility problems for blind musicians or learners wishing to read music notation. A tactile overview has been used to address problems associated with understanding the structure or ‘shape’ of a piece of music. This sets Weasel apart from other systems (electronic and non-electronic) which tend to focus on music at an event level. The user interacts with the overlay to access music information contained within the piece. This can be explored at different levels using either speech-output or audio-playback depending upon individual needs. It is anticipated that the system will be adaptable such that compositional tasks are feasible. It is also believed that there is potential for this system to used with dyslexic music-learners who experience difficulties reading and writing CMN.

Acknowledgments

Funding to assist with the continuation of the Weasel project was provided by The Nuffield Foundation (Grant Number: NAL/00477/G) as part of the new lecturers in science, engineering and mathematics scheme.

References

1. Bruce, McKennell, Walker: Blind and partially sighted adults in Britain: the RNIB survey. Her Majesty’s Stationary Office London (1991)
2. Crombie, Dijkstra, Lindsay, Schutt: Spoken Music: Enhancing Access to Music for the Print Disabled. Lecture Notes in Computer Science, Vol 2398 (2002) 667-674
3. Challis, B.P., Edwards, A.D.N.: Weasel: A System for the Non-Visual Presentation of Music Notation. In Proceedings of ICCHP 2000 (2000) 113-120
4. Challis, B.P.: Weasel: Access to Music Notation for Blind People. Cultivate Interactive, Issue 8, November 2002. http://www.cultivate-int.org/issue8/weasel/ (2002)
5. Challis, B.P.: Design Principles for Tactile Communication Within the Human Computer-Interface. PhD Thesis, University of York, UK (2000)
6. Challis, B.P.: Designing Interactive Tactile Diagrams. In Proceedings of ICCHP 2002, Lecture Notes in Computer Science, Vol. 2398. Springer-Verlag, Berlin Heidelberg New York (2002) 559-560

Math Class: An Application for Dynamic Tactile Graphics

Peter Albert

Handy Tech GmbH, Weltzienstr. 43, 76135 Karlsruhe, Germany
Peter.albert@handytech.de

Abstract. Research work has been ongoing in the field of dynamic display of tactile graphics for about forty years now. Compared to these numerous efforts, the number of systems actually in use in daily life has remained quite low. Today, there are still very few workplaces for the blind that include a dynamic tactile graphic display for providing information in graphic form, besides speech and Braille output systems. One application which does utilize such devices today, however, are math classes in schools. The following paper introduces a system that permits blind students to both create and explore mathematical graphics without assistance.

1 System Requirements

The concrete problem that set in motion the development of this system was the increasing use of pocket calculators with graphic displays in math classes in public schools. These graphic-display calculators are in use not only during daily classroom instruction, but also for tests and even in final exams. This meant that there was an urgent need to develop a system with equivalent capabilities for blind students integrated in mainstream schools.

The requirements for such a system, to enable blind pupils and students to quickly create and then explore tactile mathematical graphics, were the following:

- The blind users must be able to create the tactile graphic displays on their own.
- It must be possible to create the graphic displays in the classroom environment at speeds that are comparable to graphic calculators.
- The coordinates of points should be provided as speech output or displayed on a Braille display.
- Customizable enlargement and an exploration of the graphic in sections must be possible.
- Curves and the contours of geometrical objects must be displayed as an one-dot-wide line regardless of enlargement.
- Separate mathematical objects must be distinguishable and it must be possible to display them separately.
- Orientation aides should be available to show where the currently displayed section is located.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 1118 – 1121, 2006.

2 The Maple-GWP-System

A dynamic tactile display called GWP should be used as the output device within the new system. The GWP consists of a matrix with 16 x 24 dots working with piezoelectric actuators. There are function keys on the GWP unit for zoom, navigation, and to activate various display and orientation aides.

Fig. 1. The dynamic tactile graphic display GWP

Other system components that contribute to the system comprise the following:

- The math program Maple
- The screen reader JAWS, for operating the math program and for the output of co-ordinates
- The custom software Plotexplorer

Figure 2 shows the system components of the Maple-GWP system, and the communications channels between the components.

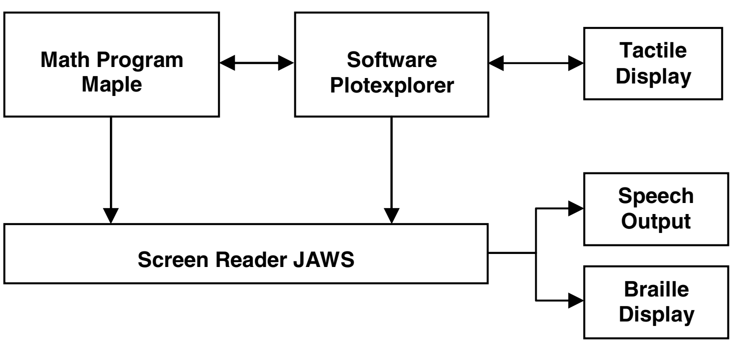


Fig. 2. System Components of the Maple-GWP System

The mathematics program Maple offers a wide range of functions, making it suitable for schools, universities, and professional use. Maple makes a range of commands available that can be used to create mathematical graphics on the screen. The data created by Maple for the screen graphics can be accessed through a defined interface by an application program.

JAWS is used as a screen reader to provide access to Maple. A communications interface from Plotexplorer to JAWS is used to have the system either announce in speech or display on an associated Braille display the coordinates of points on the graph.

The Plotexplorer software processes Maple's graphics data for display on the tactile display, and also provides a series of display aides that assist a blind user with an in-depth exploration of the mathematical graphics.

A blind student will use the Maple program in math class. If a graphic representation is needed, an entire package of plotting functions is available in Maple. An additional command can then be entered into Maple, and the graph will not only be displayed on the PC screen, but also on the tactile display.

When this command is entered for the first time, Plotexplorer starts up as a background process. Plotexplorer processes the graphics data provided by Maple for representation on the tactile display. Plotexplorer also provides the functionality needed for exploring the mathematical graphic. The following functions are available for this purpose:

Zoom and Navigation. Once a graphic representation has been output from Maple to the tactile display, it is shown just as it is shown on the PC (full view). If the user zooms in on it, an enlarged portion of the full view is represented on the tactile display. This partial view will be referred to as the section from now on. The navigation keys on the display unit can be used to move the section around. This allows a step-by-step exploration of the overall graph.

Types of Representation. Maple provides a minimum of 50 precisely calculated anchors for the graphical output. In full view mode, there are usually several such anchors associated with one tactile pin. At a powerful zoom rate, however, these anchors move away from each other until there is a discernible distance between them.

There are two types of representation, the so-called "dot mode" and the "line mode." In dot mode, the only output are the anchors, which are shown as dots by raised pins on the tactile display. In line mode, the anchors are connected by lines, so that there is a continuous line even at high zoom rates.

Blinking Cursor, Output of Coordinates. A blinking cursor (pin moving quickly up and down) indicates a dot on the tactile display that includes one or several anchors of a mathematical object. The numerical values of the central anchor indicated by the blinking cursor can be forwarded to JAWS for output. The values are announced by the speech output, and appear as pairs of numbers at the front of the Braille display line. After a certain period of time, this numbers pair will disappear. The time-out period for the Braille display can be controlled by the user. In most cases, the coordinates will be rounded off for output. The user can choose one of two powers of rounding.

Marks. In order to find a certain place within a graph again quickly, anchors can be marked. A marked location is indicated on the tactile display by a slowly blinking pin.

Navigation by Using Marks. Regardless of the enlargement rate, the user can jump to marked locations with the blinking cursor. If the next marked location happens to be outside of the currently displayed section, the section will be moved automatically until the marked location is on the edge of the image section.

Stretching of a Section. The section between the current position of the blinking cursor and the nearest marked location to the left can be stretched, so that the marked location becomes the left edge of the section, and the blinking cursor the right edge. This makes it easy to enlarge interesting parts of a curve quickly and on target.

Selecting Objects. Several mathematical graphs can be represented on the tactile display simultaneously. Plotexplorer is capable of distinguishing the various graphs. The user can therefore choose to display only any one or all graphs at once. The axes can also be turned on or off.

Orientation Mode. Orientation mode is available in case the user becomes disoriented as to where the section currently displayed on the tactile display is actually located within the graphic. By switching into orientation mode, the current section is represented as a blinking rectangle within the full view mode. This gives the user an overview as to which area of the overall graphic the current section on the display is showing.

All Plotexplorer functions can be operated via the function keys on the tactile display in full. In addition, the Plotexplorer functions are also available for sighted assistant operators via a graphic user interface. Based on this dual operator concept, an assistant and blind students can use the system together during a training phase.

Practical experiences with this system so far have shown that blind students master the functions available for exploring mathematical graphs rather quickly. The commands that are required for creating such graphs in Maple are also quickly learned. This means that students are capable of using the system on their own during classroom instruction after a relatively brief training period. Educational agencies in Germany have evaluated this system and have approved it as a resource for blind students integrated in mainstream schools.

3 Future Work

There are two aspects that can contribute to rescuing the use of dynamic tactile graphics from its current niche existence in the future, thus making the integrative effects of these systems available to a broader circle of users. First, what's needed is a new actuator technology for the tactile pins that is more compact and especially less expensive than the piezo technology in use today.

Second, it will be necessary that graphical information are represented in an open accessible standard, which provides additional descriptive information about the graphical objects. Vector graphics using the SVG language is here a promising approach.

Helping People with Visual Impairments Gain Access to Graphical Information Through Natural Language: The *iGraph* System

Leo Ferres, Avi Parush, Shelley Roberts, and Gitte Lindgaard

Human-Oriented Technology Laboratory
Carleton University
Ottawa, Ontario, Canada
{leo_ferres, avi_parush, gitte_lindgaard}@carleton.ca,
smrobert@connect.carleton.ca

Abstract. Much numerical information is visualized in graphs. However, this is a medium that is problematic for people with visual impairments. We have developed a system called iGraph which provides short verbal descriptions of the information usually depicted in graphs. This system was used as a preliminary solution that was validated through a process of User Needs Analysis (UNA). This process provided some basic data on the needs of people with visual impairments in terms of the components and the language to be used for graph comprehension and also validated our initial approach. The UNA provided important directions for the further development of iGraph particularly in terms of interactive querying of graphs.

1 Introduction

Graphs (line graphs, pie charts, etc.) are common and pervasive: they have been the accepted norm for the presentation of complex numerical data since at least the 18th century [1]. As [2,3] point out, this manner of presenting numerical data takes advantage of the human ability to recognize certain visual patterns that the numbers by themselves (e.g. in tabular form) do not readily reveal. By quickly looking at a line graph, for example, we can determine, with less precision than knowing the real numbers, where the minimum and maximum values are, whether the values have an upwards or downwards trend or whether the line is smooth or "pointy", the scale, whether it looks like an "M" or like a semicircle or like a bell, etc. All these little pieces of information (and many more we have not mentioned here for lack of space) allow us to interpret the numbers and hypothesize about their relation on many dimensions, helping us draw conclusions about the nature of the data being represented.

All the cognitive mechanisms used to find these visual patterns [4,5,6,7] depend, rather obviously, on the human visual faculty [8]. If this faculty is not present, or considerably damaged, as is the case with blind people and people with visual impairments, then the problem resides in finding alternative ways to

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 1122–1130, 2006.

make (at least some of) the visual information encoded in graphs available to this audience. To this purpose, we implemented iGraph, a system that provides short descriptions and a query system to work with graphs. However, not any kind of description will do, and, on the other hand, describing everything there is to a picture (a graph in this case) may become, at times, unfeasible (e.g. when there are one hundred data points, or more than three lines in a line graph, etc.). Thus, our problem is not only an engineering or implementation problem, but a problem of verifying the need, building a better language model and validating the solution that iGraph provides. This paper reports the outcome of a User Needs Analysis (UNA) and Wizard-of-Oz (WOZ) study towards an improved version of iGraph. We first review the state of the art in R&D and some currently available applications. Second, we introduce our first implementation of iGraph. The third and main part of the paper is dedicated to assessing the needs and a validation of the concept of iGraph. Finally, we discuss future implementations of iGraph on the basis of the findings of the UNA and WOZ studies.

2 Related Work

The idea of presenting graphs multimodally is not new. In the digital arena, there are currently two ways of helping people with visual impairments gain access to and work with the information presented by graphs: the “sonification” option and the “speech” or “reader” option. In the former, researchers have been exploring graph sonification and comprehension since at least the early 1990s [9,10,11]. Typically, in these systems, the values of the ordinate of a graph are converted to the different notes and octaves of a musical instrument and time usually stands for the abscissa of the graph.

For the NL (Natural Language) option, several studies investigating how to provide natural language descriptions of data graphs have been carried out and demo systems have been developed [12]. Worth noting are, for instance, PostGraphe and SelText, systems that generate a graph and its caption from an input of data in tabular form and some intentions from the users concerning the message the graph should convey [13,14,15]. In the same tradition of caption generation is SAGE [16], though this system aims at generating captions for 2D diagrams and maps as well as graphs. Other systems include the generation of summaries of potentially very large databases of time-series data using time-series data summarization algorithms. Examples of these systems include weather report generation systems (e.g. SUM-METEO), summarizing climatological observations (such as temperature, wind speed, humidity, etc.) [17,18,19,20,21], and others such as SUMTIME-TURBINE [22] and NEONATE [23]. Even some relatively simple commercial applications that make use of NL methods are currently available (JAWS[1], a screen reader and ChartExplainer[2], a software environment for automatically generating natural-language summaries of charts and tables, developed by CoGenTex).

[1] http://www.freedomscientific.com/fs_products/software_jaws.asp

[2] http://www.cogentex.com/products/chartex/

3 Our Initial Approach: iGraph

The technologies mentioned in the previous section have proven useful in several areas, but they also have some important drawbacks. Auditory graphs, for example, do not work very well when the input graph is complex, containing more than three lines, or when it is not time-dependent, as in pie charts. In a pure auditory graph, legends and other linguistic pieces of information are not read, and dimensions such as color, depth, etc. are lost. Another way to represent graphs is natural language. This is so for several reasons: first, humans are more attuned to this system and we require less training in individuating different signs (words, in this case, as opposed to going from a C to a C#); second, NL is representationally a very rich way to describe much of the information shown in graphical representations. In short, a picture is worth a thousand words, but, when spoken words is all we have access to, those thousand words are enough.

This being said, the NL systems described above are far from ideal for our purposes here; namely, helping blind and visually-impaired people work with graphical representations of data. Most importantly, all the NL systems mentioned above are quite inflexible: they provide a single summary or a short caption of a graph, but not the ability for users to explore the graphs. The system we are developing offers this novel ability to interact by means of a quite flexible and scalable NL dialogue. Figure 1 below shows the main components of our system and some of the implementation techniques for the different sub-systems.

Our prototype (the *inspectGraph* system or iGraph) automatically provides descriptions of graphs and allows for natural language interaction with a graph. The input to the system comprises the encoding of the necessary (and sufficient) properties of a graph in the EXtended Markup Language (XML) format (titles of the axes, values and the interval labels for the X-Axis). To allow for the description and the querying, iGraph implements three subsystems. The first takes an XML file, parses it and writes a logic version of the given graph, together

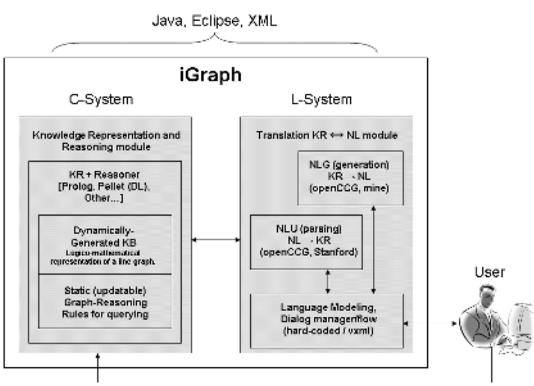


Fig. 1. The inspectGraph architecture

with several mathematical properties of the input graph (minimums and maximums, slope of the increase or decrease between two points, etc). We call this the P-System. The second subsystem stores rules for describing and querying the logico-mathematical version of the input graph generated by the P-System. We call this the C-System. The third is a language subsystem that takes the logic-like representations (and, in principle, all possible inferences) from the first and second subsystems and outputs a natural language text (both as a description and/or as the response to a query) plus the handling of a dialogue modelling algorithm for querying the graph. We call this the L-System.

As stated before, iGraph is more than an engineering problem. For the system to be useful, several issues must be considered. Two of the most important ones are *i*) what blind and visually-impaired people would ask and need to know about a graph, and *ii*) the language model needed to convey visual information successfully. To achieve this, we carried out a comprehensive UNA of people with visual impairments ranging from congenitally blind to recent blindness. The next section addresses our UNA methodology and some of the results.

4 The iGraph UNA: Method

A UNA is a family of methods used to uncover the needs and requirements a user has as well as the capabilities needed from the product to meet the user's goals. This involves understanding the target audience, their typical tasks, and their specific constraints, usually elucidated through a combination of observational techniques, including interviews, surveys and consulting with domain experts. The results provide user interface objectives, system requirements, and feature requirements[3]. The primary goal of the UNA and evaluation component of this project is to gain insight into the needs and requirements of InspectGraph users who are blind or visually impaired as well as maximize the usability of the existing InspectGraph product. This was accomplished by means of a semi-structured interview and a user-centered system evaluation.

Participants. Participants were recruited through recruitment notices posted on local mailing lists at Carleton University's Centre for Disabilities, and through personal contacts. 11 participants whose ages ran from 19 to 57 were paid $20 for approximately an hour of their time. The interviews took place in our usability lab on campus or another location that was convenient for participants (i.e. home).

Materials. Pen and paper were used to record participants' responses. A Panasonic Leica Dicomar video camera and tripod was also used to record the interview.

Design. The overall method included a) demographic questionnaire b) informal interview c) information extraction and d) a wizard-of-oz study:

[3] `http://www.usabilityfirst.com/glossary/term_576.txl`

a) **Demographics:** this questionnaire included information regarding their background in terms age, gender, education, and information regarding their blindness.

b) **Informal Interview:** a semi-structured interview was conducted to gain insight into the users previous interaction with graphs, for example, where (work, school etc), how (tactile, descriptions etc), and the type of interaction (description from others, tactile graphs, other technology etc). This interview was also used to ask participants how they would like to be able to interact with graphs and what goals they would like to accomplish if they could.

c) **Information Elicitation:** participants were told that the experimenter had a graph in front of them. They were required to ask the experimenter any questions that would help them obtain enough information in order to understand what the graph looked like. This information helped to understand the type of language these participants used so that it could be also used within the system.

d) **Wizard-of-Oz study:** The wizard-of-oz (WOZ) technique enables unimplemented technology to be evaluated by using a human to simulate the response of a system and receive immediate feedback with regards to the systems' function and usability. This technique is used to test concepts and techniques during design and provides suggestions for functionality before it is designed and implemented. In this section participants were told that they are interacting with a computer system through a natural-language interface, though in fact they are not. Instead the interaction is mediated by a human operator, the wizard, allowing the participant more freedom of expression [24]. Natural dialogue researchers conclude that empirical studies of this unique communication situation are required for the development of user-friendly interactive natural dialogue systems (Dahlbck et al., 1993). In this particular section the experimenter read out what the iGraph system would currently say to describe a graph. To investigate the participants' understanding of the spoken graph they were then asked recall questions such as "what is the highest point in the graph?"

Overall, this experiment took a freeform approach which included observations and informal questions posed to 11 participants and took approximately one hour per subject. Again, audio and video of the entire interview were recorded for the purposes of evaluation and to show the results of our testing.

Results. The demographics of the participants interviewed in terms of their vision ranged from congenitally blind to non-congenitally visually impaired. However, the differences between these groups in terms of their needs did not differ enough to look at the data separately. Therefore the data listed include all participants. Education of the participants ranged from current undergraduate (n=2), Bachelor's degree (n=8), and Graduate student (n=1). Participants' previous experience with graphs differed in terms of how long they were or were not with sight. 5 participants were sighted at least through high school, and 6 were either born blind or lost their vision prior to high school. High school was the time when all participants indicated that they had the most interaction with graphs. Overall, the informal interview revealed that all participants attempt to avoid graphs. There are no technical aids that help them extract information from graphs available at the present time, which any of them knew of. However when they do need to read graphs they rely on the support of people around them to describe the graph(s). The information extraction section of the UNA revealed the most common words used to interact with the interviewer in order to extract information from the graph. Table 1 lists the words that were most frequently used to ask questions during the information extraction section of the interview. There were no differences in terms of the words used or the requests by those who had experienced graphs visually compared to those who had not.

The wizard-of-oz study revealed that overall most participants were able to visualize what the graph looked like following the reading of iGraph descriptions. 9 of 11 participants explicitly reported that they liked how the graphs were described, and all participants reported that they could find iGraph useful. Table 2 below shows the top 5 most requested requirements. All participants requested that the purpose, type of graph, main title and axes titles are read first, followed by the general trends, and then the details. Most participants requested that they iGraph describe the data point by point, however the most common remark was that this type of description may only be useful with simple, single line graphs.

Table 1. The 14 most frequently used words used by participants to extract graph information

Case	Word	Frequency
1	X, Y Axes	41
2	line up/down	32
3	drastic/gradual shift/drop/jump	27
4	describe point by point	22
5	slope up/down	22
6	lowest/highest point	22
7	purpose of graph	21
8	range of axis	19
9	scale of axis	18
10	Main title	17
11	Type/kind of graph?	17
12	starting point	14
13	axis title	13
14	top/bottom right/left	13

Table 2. The top five requests

Case	Request	Frequency
1	Purpose, Type and Titles first (main and axes)	11
2	Describe point by point	9
3	Describe general trends and then details	8
4	Ability to repeat/pause/rewind	8
5	Keep it simple	6

5 Conclusions and Future Work

Given this UNA study, we may conclude that iGraph is indeed a useful tool. Once iGraph is finished, it should provide immediate benefits to people with visual impairments when it comes to interacting with graphs. The Wizard-of-Oz study, plus the quantitative analysis of the words and expressions used to refer to graphs allowed us to construct new, hopefully more descriptive descriptions of the input graphs. One important characteristic is that the main requirement is to provide the user with the general information first followed by the details. Below is an example of our current descriptions given the UNA and WOZ studies:

```
Description: The graph is about paper clip sales. It's a one- variable line graph.
The y axis is millions of dollars and the x axis is month. The maximum value of the
y axis is 93, the minimum is 10. The graph has a general upwards trend. Evolution:
It is quite irregular. It starts at 63. There is a moderate decrease at x1 to 23.
There is a steady increase from 23 to 91 from x1 to x3. There is a steady decrease
from 91 to 54 from x3 to x5. There is a small increase at x6 to 71. There is a
moderate decrease at x7 to 10. There is a dramatic increase at x8 to 93. There is
a moderate decrease at x9 to 38. Finally, there is a steady increase from 38 to 69
from x9 to x11.
```

These descriptions generated by iGraph, although also reporting the graph point by point, summarizes part of the graph: What the graph is about (paperclip sales), what kind of graph it is (a line graph), the title of the axes (the Y-axis is sales in millions of dollars and the X-axis is month), a general trend line given by interpolation (an upwards trend), some idea of how "steep" the lines look (small, moderate or sharp increase/decrease), minimum and maximum values and a summarization such that every time there is more than one edge increasing or decreasing, the application will output a "steady increase/decrease" from point A to point A+n: for instance, in the above example: "There is a steady decrease from 91 to 54 from x3 to x5". This provides with a more manageable load of information, even if it is at the expense of losing some of the details (how much x4 decreases wrt x3 and how much x5 decreases wrt x4, for example).

According to the studies reported in these pages, iGraph does indeed help blind and visually impaired individuals with their interaction with graphs. This will hopefully build independence and a less marked aversion to graphs as an important data exploration tool.

Acknowledgements

This paper was written with the support of HTX (Health Technology Exchange), Cognos Inc., the Ontario Research Network for E-Commerce and The Cybercartography and the New Economy project funded by the Social Sciences and Humanities Research Council (SSHRC) of Canada under the Initiative on the New Economy (INE) Collaborative Research Initiative Grant. The authors would also like to thank the members of the HOTLab, Yandu Oppacher and Zhihong Li for expert advice on coding parts of iGraph, Jing Liu for constant support and, especially, the participants who took part in the UNA and WOZ studies.

References

1. Shah, P., Carpenter, P.A.: Conceptual limitations in comprehending line graphs. Journal of Experimental Psychology: General **124** (1995) 43–61
2. Bertin, J.: Semiology of Graphics. The University of Wisconsin Press, Madison, WI (1983)
3. Carpenter, P., Shah, P.: A model of the perceptual and conceptual processes in graph comprehension. Journal of Experimental Psychology: Applied, **4** (1998) 75–100
4. Pinker, S.: A theory of graph comprehension. In Freedle, R., ed.: Artificial intelligence and the future of testing. Lawrence Erlbaum, Hillsdale, NJ (1990) 73–126
5. Trickett, S.B., Ratwani, R.M., Trafton, J.G.: Real-world graph comprehension: High-level questions, complex graphs, and spatial cognition. (under review)
6. Kosslyn, S.: Understanding charts and graphs. Applied Cognitive Psychology **3** (1989) 185–226
7. Lohse, J.: A cognitive model for the perception and understanding of graphs. In: Proceedings of the SIGCHI conference on Human factors in computing systems, ACM Press (1991) 137–144
8. Foster, M.: Evaluating models of visual comprehension. In: Proceedings of EuroCogsci 2003. (2003)
9. Flowers, J.H., Hauer, T.A.: The ear's versus the eye's potential to assess characteristics of numeric data: Are we too visuocentric? Behavior Research Methods, Instruments, & Computers **24** (1992) 258–264
10. Flowers, J.H.: Thirteen years of reflection on auditory graphing: Promises, pitfalls, and potential new directions. In: Proceedings of ICAD 05-Eleventh Meeting of the International Conference on Auditory Display. (2005)
11. Brown, L., Brewster, S., Ramloll, R., Burton, M., Riedel, B.: Design guidelines for audio presentation of graphs and tables. In: Proceedings of ICAD 2003 Workshop on Auditory Displays In Assistive Technologies, University of Boston (2003)
12. McCoy, K.F., Carberry, M.S., Roper, T., Green, N.: Towards generating textual summaries of graphs. In: Proceedings of the 1st International Conference on Universal Access in Human-Computer Interaction, ACM Press (2001)
13. Fasciano, M., Lapalme, G.: Postgraphe: a system for the generation of statistical graphics and text. In: Proceedings of the 8th International Workshop on Natural Language Generation (INLG-96). (1996)
14. Corio, M., Lapalme, G.: Generation of texts for information graphics. In Saint-Dizier, P., ed.: Proceedings of the 7th. European Workshop on Natural Language Generation (EWNLG'99), Toulouse (1999) 49–58
15. Fasciano, M., Lapalme, G.: Intentions in the coordinated generation of graphics and text from tabular data. Knowl. Inf. Syst. **2** (2000) 310–339
16. Mittal, V.O., Carenini, G., Moore, J.D., Roth, S.: Describing complex charts in natural language: a caption generation system. Computational Linguistics **24** (1998) 431–467
17. Kerpedjiev, S.M.: Automatic generation of multimodal weather reports from datasets. In: Proceedings of the third conference on Applied natural language processing, Association for Computational Linguistics (1992) 48–55
18. Goldberg, E., Driedger, N., Kittredge, R.: Using natural-language processing to produce weather forecasts. IEEE Expert **9** (1994) 45–53
19. Reiter, E., Dale, R.: Building applied natural language generation systems. Journal of Natural Language Engineering **3** (1997) 57–87

20. Reiter, E., Dale, R.: Building natural language generation systems. Cambridge University Press (2000)
21. Yu, J., Reiter, E., Hunter, J., Sripada, S.: A new architecture for summarising time series data. In: Proceedings of INLG-04 Poster Session. (2004) 47–50
22. Yu, J., Reiter, E., Hunter, J., Sripada, S.: Sumtime-turbine: A knowledge-based system to communicate time series data in the gas turbine domain. In: Proceedings of IEA/AIE-2003. (2003) 379–384
23. Sripada, S., Reiter, E., Hunter, J., Yu, J.: Summarising neonatal time series data. In: Proceedings of 2003 Conference of the European Chapter of the Association for Computation Linguistics. Companion Volume. (2003) 167–170
24. Dahlbäck, N., Jönsson, A., Ahrenberg, L.: Wizard of oz studies – why and how. Knowledge-Based Systems **6** (1993)

Fingertip Guiding Manipulator: Haptic Graphic Display for Mental Image Creation

Yoshihiko Nomura, Yuki Yagi, Tokuhiro Sugiura, Hirokazu Matsui, and Norihiko Kato

Graduate School of Engineering, Mie University, Tsu, Mie 514-8507, Japan
nomura@mach.mie-u.ac.jp
http://www.robot.mach.mie-u.ac.jp

Abstract. A fingertip guiding manipulator was developed as a haptic graphic display for blind people: it helps them create mental images of raised graphics and reliefs with stroke order. Through fingertip kinesthetic sense, while foreseeing the direction of the upcoming stroke by perceiving the knob's direction, users experience the uplifted and fallen strokes by perceiving their fingertips being pulled in the horizontal and vertical directions. Especially, the up-lifting function expands a representation capability from "*single stroke graphics*" to "*multiple stroke graphics*.

1 Introduction

Imaging mental maps in their minds, blind people usually walk by themselves usually with the help of guide sticks to explore obstacles. Therefore, they would feel uneasiness and insecure when walking unless they had previously made up mental maps, i.e., a kind of mental images. Besides the mental maps for walking, there are many situations where the mental images are used. Thus, a support system to create mental images must become a useful tool for blind persons to get higher level of communication, and to enjoy quality of life. To meet the needs, there have been reported many devices and studies in the field of mental image creation such as those by audition [1], by tactile sensation [2,3], and by both audition and tactile sensation [4,5,6]. Although these tactile systems showed effectiveness to some extent, they had cost-expensive problem. In addition, some promising haptic devices have also been proposed: a two-handed multi-fingers string-based device was used for simple cylindrical object size recognition [13], and a multi-fingered haptic interface was used for simple spherical object grasping operation [14]. There have also been some studies on haptic interfaces [15,16], the objective of which is to provide the feeling of the pencil touch motion, and is very similar to this work except that their system based on force feedback by using a highly expensive commercially available haptic device.

In place of using tactile sense, the authors have tried to develop some interactive sound-based color image pattern presentation systems as a man-machine interface: a system using a pen-tablet [7,8], and that using a touch-panel display [9]. Since those

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 1131 – 1138, 2006.

by audition can use market-produced devices, they have advantages in the availability and in cost. Regretfully, however, their image-perception performance had been substantially low.

Then, the authors had introduced a specially designed 3-DOF (degree of freedom) manipulator as a haptic graphic display [10,11]: the mechanical geometry calibration procedures and the basic perception principle were described in [10], and the image-perception performance was considerably improved as in [11]. The 3-DOF manipulator had a novel function to present line drawings. The person pinches the knob by his/her fingertips. Then, through his/her fingertip kinesthetic sense, while foreseeing the direction of the upcoming stroke based on the knob's direction, he/she practically experiences the strokes by perceiving the knob's horizontal positions. Although the 3-DOF manipulator showed an image-perception performance being comparable to the tactile graphics by the raise writer, it had a restriction that the figures to be shown should be "*single stroke* graphics".

Adding one degree of freedom by mounting an up-and-down actuator, the 3-DOF manipulator evolves into a 4-DOF manipulator. The 4-DOF manipulator is expected to enhance the mental image creating performance since it is able to represent raised graphics and reliefs with stroke order. The most prominent advantage of the 4-DOF manipulator is in a point that the up-and-down actuator expands representation capability from "*single stroke* graphics" to "*multiple* stroke graphics" by resembling the up-and-down motions to pen's up-and-down. Some experimental results are described to prove the effectiveness of the proposed system in this paper.

2 System Description

The motivation of the proposed mental-image-creation supporting system is as follows. If we taught blind persons the shapes of graphics, we would guide their hands with our own hands. However, the teaching processes will take a lot of time and efforts for both of the helpers and the blind persons. If a manipulator took the helpers place, the blind persons could learn the graphics by themselves. (See Fig. 1 & Fig. 2)

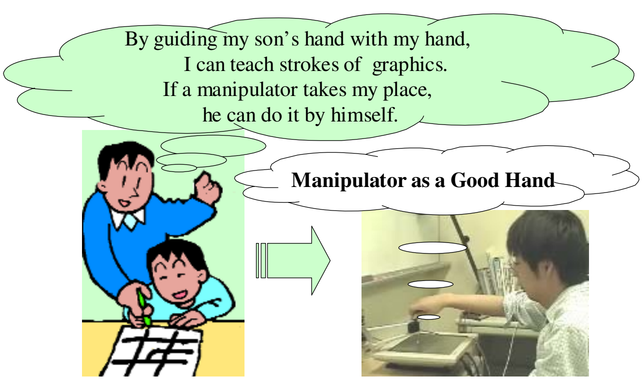


Fig. 1. Why did we employ a manipulator?

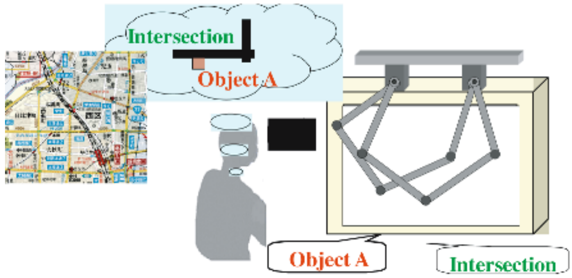


Fig. 2. Main concept of the proposed system. To create mental images of 2-D graphics such as maps and layouts, manipulator teaches object's shapes by guiding subject's fingertip, and speaker teaches object's name.

The proposed fingertip-guiding manipulator features four degrees of freedom, and help blind people create mental images of raised graphics and reliefs with stroke order. A person is assumed to grip the hexagonal prism shaped knob in use. The knob is attached to the arm's end of the fingertip guiding manipulator. Within the 17inch-display area, the horizontal position of the knob is controlled by a 2-DOF quadrilateral link that are driven by a couple of servomotors. The vertical position of the knob is also controlled within the 14mm length by a 1-DOF up-and-down actuator that is driven by a servomotor via cam mechanism. (See Fig. 3). In addition, the knob rotation angle is controlled by a rotational actuator that is driven by a stepping motor. Taking the benefit of higher traction force, the 2-DOF quadrilateral link actuator was introduced as a haptic human-computer interface in this work, and had also been introduced for upper-limb rehabilitation exercises [12].

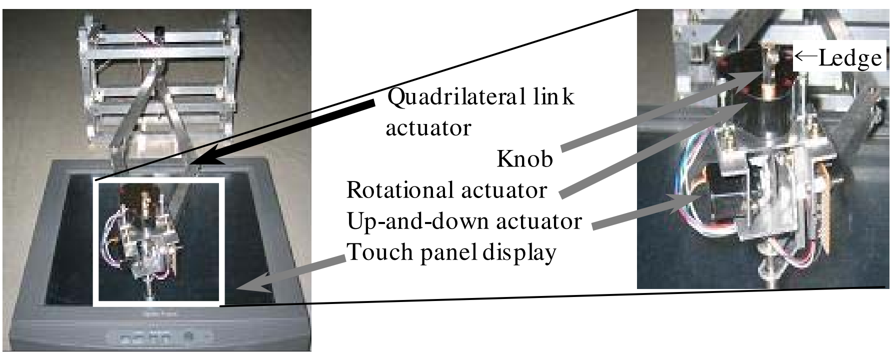


Fig. 3. The proposed haptic graphic display: the fingertip-guiding manipulator

The fingertip guiding manipulator has the haptic graphic displaying functions:.

– Indicating function: The person perceives the knob orientation via his/her fingertip tactile sensation, and predicts the directions of the upcoming strokes: the knob orientation is represented by the orientation of a hemispherical ledge, and it always indicates the directions of upcoming strokes at corners, and the tangential orientation of the currently tracing point on the stroke.

- Leading function: The person's fingertips are pulled via the knob tracing the strokes of the presented graphics, sequentially, and he/she experiences the position and shapes of the strokes through his/her arm kinesthetic sensations.
- Up-lifting function: The knob-gripping fingertips are uplifted or fallen, and he/she perceives the motion through his/her wrist's kinesthetic sensations. Then, the person experiences the pen up-and-down motions when stationary and the upheaval or caving motions when tracing strokes.

Thus, the person is assumed to perceive raised graphics with stroke order as in Fig. 4

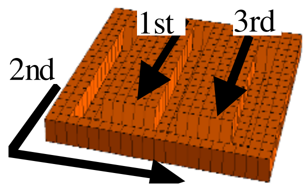


Fig. 4. An example of the relief-like presented graphic

3 Haptic Graphic Perception Experimental

The proposed haptic graphic display, the fingertip-guiding manipulator, was compared with the raise writer from the viewpoint of the performance on mental-image creation of a line drawing graphic. Fig. 5 shows a scene of the experiments using the fingertip-guiding manipulator and the tactile graphics drawn by the raise writer: the former is called the fingertip-guiding manipulator mode, and the latter the raise writer mode in this paper. In the raise writer mode, the tactile sense on finger cushions would be utilized to perceive the directions of raised line strokes on the tactile graphics, and the haptic sense, i.e., the kinesthetic sense of the upper/lower arm would be utilized to perceive the horizontal positions of the raised line strokes.

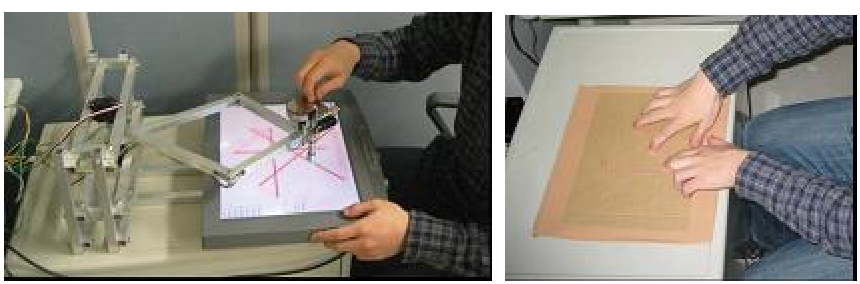


Fig. 5. Experimental setup (Left: fingertip-guiding manipulator. Right: tactile graphic)

3.1 Experimental Conditions

Description of Raise Writer Mode. In tactile graphics, lines were drawn (raised) by the raise writer, and one to four dots in raised line square 1.5 cm on a side were also drawn (raised) at the starting point on each of the strokes so as to represent the stroke order from first to fourth.

Description of Fingertip-Guiding Manipulator Mode. In contrast to this, in the case of the fingertip-guiding manipulator, the stroke order is presented through a synthesized speech. The system phonates Japanese words to inform the stroke order, staying at the starting points: for example, “3 番線スタート” that means “The third stroke starts” in English. To inform the intersecting stroke’s number, it phonates another Japanese words, staying at the intersection: for example, “2 番線交差” that means “It intersects with the second stroke” in English. The velocity of manipulator in the horizontal direction was set at about 60 mm/sec, and it took approximately 60 sec for each graphics including rotation when undergoing stroke changes. That in the vertical direction was set at about 50 mm/sec.

Presented Graphics. Four kinds of irregularly shaped multi-stroke line drawings were employed as graphics. Every graphic is composed of the same kinds of four lines so as to equalize geometrical complexity as much as possible, but differs from the others in a topological manner, as shown in Fig. 6: a vertical straight line segment of 17 cm in length, an upward-sloping one of 20 cm, a horizontal one of 17 cm, and a downward-sloping one of 20 cm. The graphics presented for the raise writer mode were also presented in the fingertip-guiding manipulator mode, changing their orientation by 45 degrees.

Experimental Sessions and Subjects. The comparative perception experiments were consecutively carried out for all the four kinds of graphics in either the raise writer mode or the fingertip-guiding manipulator mode. The four graphics experiments for each of the two modes formed a session. The order of the raise writer session and the fingertip-guiding manipulator session was counterbalanced between the subjects. The subjects were novice blindfolded 10 sighted persons of 22-24 years old. Before tests, subjects practiced adequate training to get used to both modes. At tests, subjects were allowed to exploit graphics with no time limit until they content themselves with the perception of the graphics.

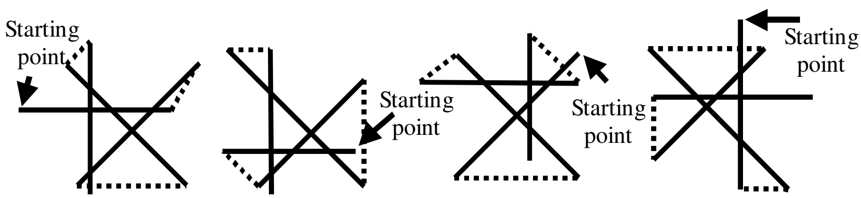


Fig. 6. Presented figures (Solid Line: actual stroke Broken line: inter-stroke motion)

3.2 Evaluating Method

The performance was measured by the consumption time to exploit the presented graphics, and the correct answer rates (CAR). CAR is defined as the ratio of the number of correctly perceived items to the total number of presented items. The items are the order, the starting point, and the topological position of each stroke, and, therefore, the total number of presented items are given as 12(=4×3).

$$\text{CAR} = \frac{\text{The number of correctly perceived items}}{\text{The total number of presented items}} . \tag{1}$$

3.3 Experimental Results, Findings, and Discussion

Fig. 7(a) shows an experimental result on the consumption time. Although the consumption time scattered in wide range for both modes, the average consumption time for the fingertip-guiding manipulator was 111.8 seconds and was much less than that of 224.7 seconds for the raise writer. The one-tailed test involving Student's *t* distribution, and not assuming homoscedasticity was carried out with the difference of the means. The test statistic *T* of the two samples was given as 3.9, and it is very much greater than a critical value of 1.68 for the significant level of 5%. Therefore, we conclude that the fingertip-guiding manipulator mode was superior to the raise writer mode from the viewpoint of the consumption time.

Fig. 7(b) shows an experimental result on CAR, i.e., the correct answer rate. The CARs for both modes achieve very high performance, and shows rather narrower scatters than the consumption time shows. It is because subjects were allowed to exploit with no time limit in this study. However, the average CAR for the fingertip-guiding manipulator was 99 % and was superior to that of 97% for the raise writer. Much more attention and effort might be consumed in active search process in the raise writer mode, and, therefore, the consumption caused the memory retention loss of the perceived graphic patterns. In contrast, in the fingertip guiding manipulator

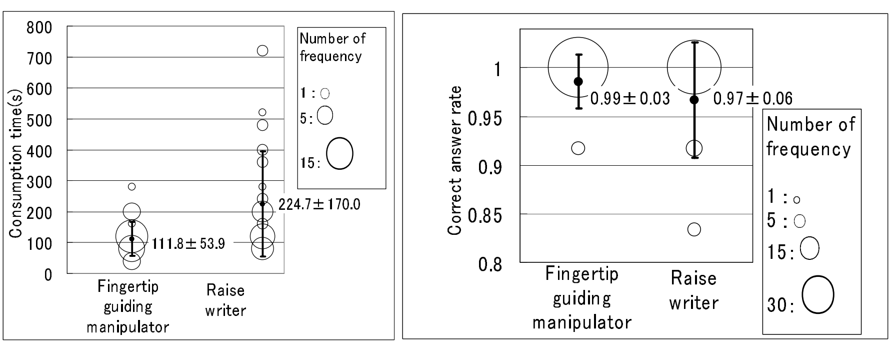


Fig. 7. Experimental results (Left: consumption times. Right: correct answer rate.)

mode, the subjects might be able to devote their processing resource into the mental image creation and the retention of the mental images. The one-tailed test involving Student's t distribution, and not assuming homoscedasticity was also carried out with the difference of the means. The test statistic T of the two samples was given as 2.3, and it is very much greater than a critical value of 1.67 for the significant level of 5%. Therefore, we conclude that the fingertip-guiding manipulator mode was also superior to the raise writer mode from the viewpoint of the correct answer rate.

4 Conclusion

A novel haptic graphic display, i.e., a fingertip guiding manipulator was developed for blind people. It helps human create mental images of raised graphics and reliefs with stroke order. The results are summarized as follows:

1. A 4-DOF fingertip guiding manipulator was developed. It is able to teach graphics with stroke order.
2. The position of the knob traces the strokes of the graphics, sequentially, and the orientation of the knob indicates the direction of the upcoming stroke, and the tangential direction at the point on the currently tracing stroke. Thus, the indicating function and the leading function are realized.
3. Mental image perception performance of the proposed system was compared with the raise writer. As a result, it was confirmed that there are prospects of practicability for the proposed system.

In the future, the authors are planning to carry out some experiments with blind people. Furthermore, there can be other way of utilizing the fingertip guiding manipulator. For example, we may utilize the manipulator not only in the passive way that the fingertips of the subjects were guided by the manipulator as in this paper, but also in the other active way that subjects actively moves the manipulator to the place where they would like to exploit. As a result, it is expected to enhance the graphic perception performance.

Acknowledgements

The authors would like to express their appreciation to Mr. Mikio Ueda, Mr. Michio Kado, and Mr. Shigeki Morisaki for their helpful cooperation and advices as blind persons. This research is supported financially by The Ministry of Education and Science, Japan: Grant–in-Aid for Scientific Research (B) 17300184.

References

1. Beddoes, M. P., Suen, C. Y.: Evaluation and a Method of Presentation of the Sound Output From the Lexiphone -A Reading Machine for the Blind. IEEE TRANSACTIONS ON BIO-MEDICAL, Vol.BME-18, No.2 (1971) 85-91
2. Simizu, Y., Saida, S., Shimura, H.: Tactile pattern recognition by graphic display: Importance of 3-D information for haptic perception of familiar objects. Perception & Psychophysics, 53 (1993) 43-48

3. Tang, H., Beebe, D. J.: A Microfabricated Electrostatic Haptic Display for Persons with Visual Impairments. IEEE TRANSACTIONS ON REHABILITATION ENGINEERING, Vol.6, No.3 (1998) 241-248
4. Shimojo, M., Shinohara, M., Fukui, Y.: Human shape recognition performance for 3D tactile display. IEEE Transactions on Systems, Man, and Cybernetics, Part A 29(6) (1999)637-644
5. Kobayashi, M., Watanabe, T.: Communication System for the Blind Using Tactile Displays and Ultrasonic Pens-MIMIZU. Proceedings of 9th Annual International Conference ICCHP 2004 (Computers Helping People with Special Needs), Vol. LNCS3118 (2004) 731-738
6. Blenkhorn, P., Evans, D. G.: Using speech and touch to enable blind people to access schematic diagrams", Journal of Network and Computer Applications. 21 (1998) 17-29
7. Nomura, Y., Nishihama, S., Ozaki, M., Sugiura, T.: Image Pattern Perception by Interactive Sound Presentation. Proceedings of 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (2001) CD-ROM Paper#369
8. Nomura, Y., Nishihama, S., Sugiura, T., Matsui, H., and Kato, N.: Auditory Sensation Aided Image Perception: Integration of Somesthetic Sense of Fingertip Position, Proceedings of 8th Annual International Conference ICCHP 2002, LNCS2398 (2002)448-454
9. Nomura, Y., Kakehashi, H., Sugiura, T., and Kato, N.: Mental Map Creation Support System for Blind Person –Fusion of Verbal Message and Somesthec Sense, Proceedings of HCI International 2003, Vol. 48, (2003) 1315-1319
10. Nomura, Y., Kakehashi, H., Sugiura, T., Matsui, H., and Kato, N.: An Indicating, Leading Manipulator as a Good Hand at Teaching Strokes: A Mental-Image-Creation Support System. Proceedings of 9th Annual International Conference ICCHP 2004 (Computers Helping People with Special Needs), Vol. LNCS3118, (2004)705-712
11. Nomura, Y., Omoto, M., Sugiura, T., Kato, N., and Matsui, H.: Direction-Indicating, Traction Manipulator to Support Mental Image Creation of Line Drawings, CD-ROM Proceedings of HCI International 2005, (2005) 1-6
12. Takanashi, Y., Terada, T., Inoue, K., Ito, Y., Lee, H., and Komeda, T.: Upper-Limb Rehabilitation Exercise Using Haptic Device System. International Journal of Human-friendly Welfare Robotic Systems, Vol.4, No.2(2003)18-22
13. Walairacht, S., Ishii, M., Koike, Y., and Sato, M.: Two-Handed Multi-Fingers String-Based Haptic Interface. IEICE Trans. Inf.&Syst.,Vol. E84-D, No.3(2001)
14. Kawasaki, H., Hori, T., Mouri, T.,:Multi-Fingered Haptic Interface Opposed to Human Hand. J. Robotics Society of Japan,Vol. 23 No. 4, (2005) 449-456 (in Japanese)
15. Inami,N., Tomonaga, H., Matsubara, Y., and Yamasaki, Y.: Pencil Touch Learning System for Japanese Handwriting Based on the Haptic Device. Trans. Institute of Electronics, Information and Communication Engineers. D-I, Vol.J87-D-I(12) (2004.12)1128-1135(in Japanese)
16. Nikolakis, G., Moustakas, K., Tzovaras, D. and Strintzis, M.G.: Haptic representation of images for the blind and the visually impaired, CD-ROM Proc. HCI International 2005

Graphic Editor for Visually Impaired Users

Atsushi Nishi[1] and Ryoji Fukuda[2]

[1] Kyushu Institute of Technology, 680-4 Kawazu 820-8502 Iizuka, Japan
[2] Faculty of Engineering Oita University, 700 Dannoharu 870-1192 Oita, Japan

Abstract. Additional adequate graphical contents are very effective in various communications. The same is clearly true of information given by a visually impaired person. The main purpose of our system is to create graphical contents without any visual information. The targets are fundamental mathematical graphic objects. Making use of a tactile pin display and ultrasonic pen, users input graphical object hearing several audio assists.

1 Introduction

Graphical contents often make the complex easy to understand. On the other hand, it would be very difficult for visually impaired persons to create graphical contents. Our main purpose is to provide a tool for creating graphical contents without visual information.

Assistive software using a tactile display or audio outputs getting popular in research. For example, Mimizu system [1] by M. Kobayashi and T. Watanabe is a drawing tool for tactile display and a system for geometric education by S. Rouzier, B Hennion, T. P Segvia, D. Chêne [2] is based on audio outputs and a tactile display. We use a tactile pin display “DotView” by KGS Corporation and an ultrasonic pen by Pegasus Technologies Ltd.. MIMIZU system was demonstrated in the last conference (ICCHP ’04). The hardware of our system is same with Mimizu system used in the demonstration.

A graphical content often includes huge information; however, we usually get necessary information at a glance. Eyes (or a brain) may pick it up very quickly. This ability is a large barrier when visual information can not be available. Most mathematical graphical contents consist of simple lines or curves, and their positional relations express basic properties in the necessary information. Our graphic editor enables totally blind user to edit mathematical graphical objects: straight lines, parabola curves, circles, ellipses and sine curves, using a tactile display and an ultrasonic pen. Some basic explanations are read out by mechanized voice. These are positional properties of one line object or positional relations of two line objects.

2 Outline of the System

Our system is a graphical editor for two-dimensional graphs. The target objects are straight lines, circles, ellipses, parabolas and sine curves. At the starting

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 1139–1146, 2006.

point of these research, we limit the targets only five object types and try to find out essential difficulties for sightless creation of graphical contents.

During use of our system, there are three phases: *Entry*, *Edit* and *Move*. We use a tactile display *DotView* and an ultrasonic pen. The tactile display consists of a pin display area (32 x 24), seven keys (four direction keys and First Finger Key, Status Key and Center Key) and a joystick (Fig 1). Through these three phases, a user edits graphical objects by the joystick and the three keys. Putting a finger on the pin display the user check the shapes or positions; putting the pen on the pin display the user hear some voice explanations. In the first phase *Entry*, five mnemonic symbols are displayed on the left edge and their right side there listed several entry marks(Fig. 2). *Entry mark* is a sign for input graphic objects. One entry mark consists of 2 x 2 pins, and these are flat at beginning. Using the joystick and the keys, a user adds a new graphic object, then these pins are raised up (Fig. 2).

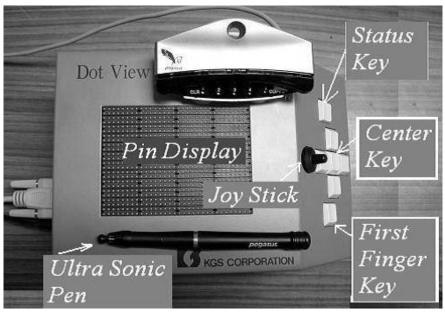


Fig. 1. Tactile Display

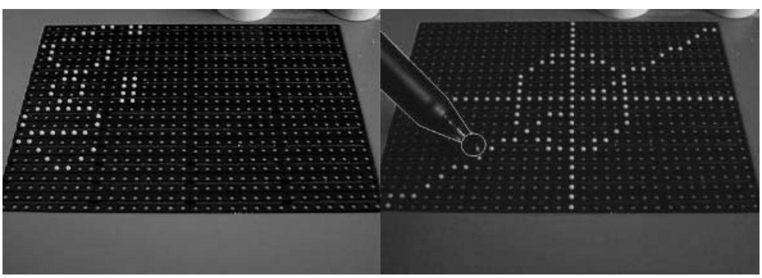


Fig. 2. Entry and Edit Phases

Entry Mark are blinking when the user input a graphic object. This blinking stands for the mode *Changing*, that is, this object changes its shape and position in the phase *Edit*.

During the mode *Changing*, the blinking stops when the user push *Center Key* one more time. Then the mode changes to *Display*, and this object are displayed in the phases *Edit* or *Move*.

After one more push of *Cener Key*, the mode of the object is *Hidden*. Moreover, the mode returns to the first mode after one more push.

The second phase is *Edit*. There are only one object which mode is *Change*, and several objects which mode is *Move*. A user changes the object of *Change* with its shape or position adjusting the relations among *Move* objects. There are three types *Shift*, *Rotate* and *Scale* of *Move*. The user select one of them using *First Finger Key* and *Status Key*.

In the last phase *Move*, the whole plane are shifted or zoomed up or down. In this phase, all parameters of each objects do not change.

3 Voice Explanation

If a user creates graphical contents without any explanation, the user must memorize all graphical objects and their positional correlations. Since the system are controlled by only three keys and a joy-stick, there are several phases in each situation, he must also memorize these temporal phases. Therefore some remainder must be useful. The following explanations are given by voice.

1. Explanation for *Entry*.
2. Explanation for one object.
3. Explanation for two objects .

In *Entry* phase, five mnemonic symbols and entry marks are listed in the pin display. When the user pushes an entry mark by the pen, corresponding explanation is read out. In *Edit* or *Move* phases, the explanation with respect to a curve is read out when the user pushes on the curve by the pen.

We will explain details of their explanation in the following subsections.

3.1 Explanation for *Entry*

Simple explanation for *Entry* is given according to the following format.

(c̃urve name) with the (*reference point name*) is (*position name*).
(position name) is one of { first quadrant ∼ fourth quadrant, on the $x(y)$-axis with $x(y)$ coordinate is positive(negative), around the origin }(Fig. 3). For a curve name of a sine curve, most adequate one is selected among four combination plus(this is not read out) or minus and sine or co-sine, according to the phase angle difference.

Table 1. Curve Name and Reference Point Name

Curve name	Reference point name
a straight line	first reference point
a parabola with convex downward(upward)	peak point
a circle	center
an ellipse	center
a (minus) sine (cosine) curve*	starting point

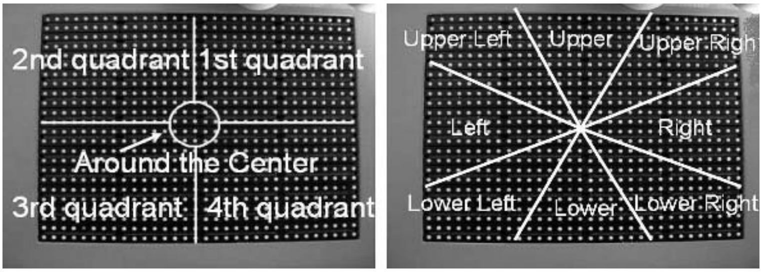


Fig. 3. Explanation For Position And Direction

3.2 Explanation for One Object

In the phase *Edit* or *Move*, there are two explanation modes: “Individual Explanations” and “Relations Between Two Objects”. These are interchanged with each other by tapping at left beside of the pin display. Each explanation mode has two sub modes: “Positional Explanation and Parameter Explanationff and “Cross Relation and Positional Relationff. These are exchanged with each other by tapping at right beside of the pin display (Fig. 4 at the end of this section).

Table 2. Positional Explanation

Straight line	The gradient and the y-intercept
Parabola	The peak point and the coefficient for x^2
Circle	The center and the radius length
Ellipse	The center and the long and short axis lengths
Sine Curve	Adequate curve name and oscillation

Table 3. Parameter Explanation

Straight line	The formula as a function
Parabola	The formula as a function
Circle	The center and the radius length
Ellipse	The center and the long and short axis lengths
Sine curve	The formula as a function

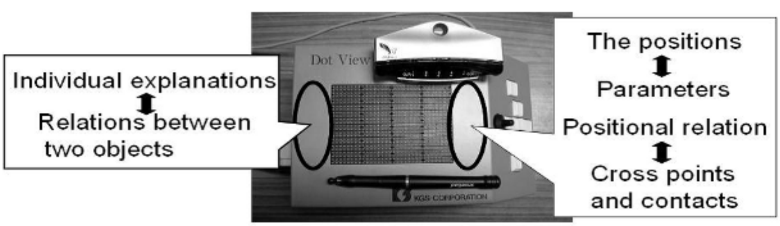


Fig. 4. Tapping Area for Exchange

Details for Positional Explanation and Parameter Explanation are listed in Table 3.2. We have no stereo type explanation for them, and to find some simple generalization of these explanation is our future problem.

3.3 Explanation for Two Objects

In the case where there are two objects in the display area, we have two explanation type: *Cross Relation* and *Positional Relation.* In Cross Relation mode, number of cross points and contact points are read out. In Positional Relation mode, direction between two reference points are read out. Moreover, when one of target objects is a straight line or a circle and two objects have no cross point and contact points, the direction of the move by which the two objects touch to each other is read out.

In general, cross points and contact points are obtained by checking all combinations of finely divided two sequence of line segments. Some special cases are "Straight Line and Straight Line", "Straight Line and Parabola", "Straight Line and Circle" and "Circle and Circle". In these cases, we are able to get mathematical solution directly, for cross points and contact points. Then we can get all cross points and contact points. On the other hands, using general method, these depend on the range of parameters. Straight Line, Parabola Curve and Sine Curve are graphs of functions. Then the searching area depends on the range of x, that is, this depends on the size of pin display area.

4 How to Use the System

In this section, we explain some details of the system, according to an actual example. Suppose that the objective graph satisfies:

1. There are a circle and a straight line.
2. The line crosses with x-axis at $x > 0$ and, with y-axis at $y > 0$.
3. The circle does not cross with the line and the circle center is in third quadrant.

We make these graphical contents as follows.

Starting with in the *Entry* phase, add two objects: a circle and a straight line. If the entry mark of the circle is blinking, move the cursor to the entry mark of a circle and push Center Key. Then the entry mark stops blinking (Fig. 2). If the entry mark of the straight line is not blinking, push the center key several times until blinking of the mark. Then the straight line is the target for changing shape.

Next, change the phase to *Edit* by pushing the first finger key. Since the first mode of this phase is *Shift*, by Status Key choose the mode *Rotation.* Then, adjust the angle of the line. Shift the line and, if necessary, check the position using voice explanations.

Return to the phase *Entry* and set the entry mark of the circle blinking. Change the phase to *Edit*, and shift the circle until the circle does not crosses or touches with the line. Then the message "No cross points and contact points" are read out when tapping on the circle by the pen (Fig. 5).

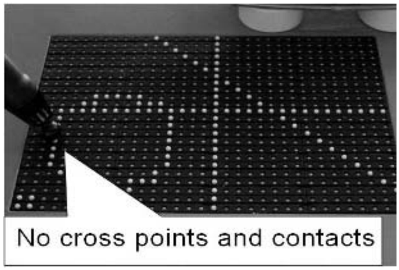


Fig. 5. Input Circle and Straight Line

5 Contact Move

The condition that two curves touch each other is very delicate with respect to moves. However, this condition often implies important aspect from a mathematical point of view. Then the following functions are necessary.

F-1. To shift or to scale up (or down) an object to touch to another curve.
F-2. To shift or to scale up (or down) an object under keeping to touch to another curve.

The target object types are Straight Line, Parabola Curve and Circle. The type of the object which the target object touches to is not limited. Considering with the further extension, the necessary conditions are:

C-1. The curve are expressed as a one parameter vector function.
C-2. For an arbitrary parameter, corresponding point is given.
C-3. The direction of the curve at any point is given.
C-4. The curve is a graph of some function or a closed curve, and when it is a closed curve, inside or outside must be determined for each point.
C-5. For an arbitrary point, corresponding parameter is given.

We can construct the function F-1 and F-2 using the condition C-1 $\sim$ C-5. We will explain about F-1 and F-2 for Parabola Curve, and other cases are similar to it.

F-1 and F-2 calls the following function.

F-3. To make a touching parabola curve at any point of a curve, when a coefficient of x^2 is given.

The differential coefficient of the parabola function is determined by the direction vector (C-3). An exceptional case is that the direction is vertical, there is no parabola curve satisfying the condition. Otherwise the parabola curve is uniquely determined, if a coefficient of x^2 and an x-coordinate of the contact point are given.

For the function F-1, finding the nearest point in a given curve with respect to the original parabola curve, the desired parabola curve is given by F-3.

For the function F-2, finding the parameter of a given curve corresponding to the contact point, and changing this parameter (for shift) or the coefficient of

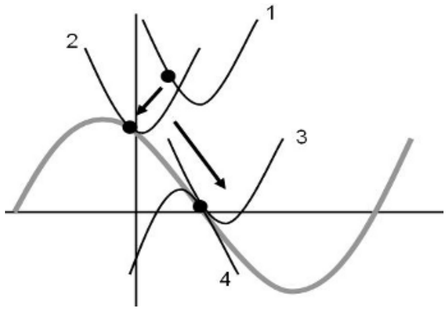


Fig. 6. Contact Move of Parabola

x^2 (for scale up or down), we obtain the desired curve as a touching parabola curve by F-3.

In Fig. 6, **1** is moved to **2** by F-1 and **2** is moved to **3** by F-2(shift) and **3** is moved to **4** by F-2(with parameter change).

6 Evaluation

To evaluate our system, we asked five persons to use the system. All of them read Braille on a daily basis, and are not accustomed to mathematics or graphics. Their details are listed below. For all of them, we gave a same test: Starting with, the

Table 4. Personal Details

	Age	Sex	PC use
A	51	male	sometimes
B	49	male	every day
C	51	male	very often
D	55	male	every day
E	55	male	sometimes

Table 5. Personal Remarks

	Positive Remarks	Negative Remarks
A	Entry phase is OK. Unit display is OK. Move is easy to understand.	Cross is difficult to understand.
B	Entry phase is OK.	Can not recognize Mnemonic (Especially for Ellipse)
C		Can not recognize the shape. Axis disturbed the recognition.
D	Entry phase is OK. Move is good. Shapes are imaginable.	Can not recognize Fast moves. Difficult to use.
E	Move is good.	Can not understand details

collaborator (test subject) input a circle and a straight line. Next, he touches the pin display and checks that the line crosses the circle. Then, he moves the line until the line does not cross the circle, and the last step is to move it until the line crosses again. The following is a list of remarks after (or during) uses of the system.

7 Conclusion

We made a system to create graphical objects without visual information. Totally blind user could input some graphic objects using voice explanations. Every one pointed out roughness of the curves, however, they may be adjust them. Indeed, one of them also left comments about this capability. Considering these facts, the following is a list of our future problems.

1. Evaluation test for long time use.
2. More effective reminder for each situation.
3. Adding many other object types.
4. Three dimensional Graphics and Other.
5. Link with other documents.

References

1. Kobayashi M., Watanabe T.: Communication System for the Blind Using Tactile Displays and Ultrasonic Pens-MIMIZU. ICCHP 2004 , LNCS 3118, pp. 725-730, 2004
2. Rouzier S., Hennion B., Martin Rotard M., Otte K., Segovia T. P., Chêne D. Touching Geometry for Blind Pupils ICCHP 2004, LNCS 3118, pp. 697-704, 2004
3. Shimojo M., Shinohara M., Tanii M., Shimizu Y.: An Approach for Direct Manipulation by Tactile Modality for Blind Computer Users :Princeple and Practice of Detectiong Information Generated by Touch Action, ICCHP 2004 , LNCS 3118, pp. 739-744, 2004

Exploiting SenseCam for Helping the Blind in Business Negotiations

Shuaib Karim, Amin Andjomshoaa, and A. Min Tjoa

Vienna University of Technology
{skarim, andjo, amin}@ifs.tuwien.ac.at

Abstract. During business meetings, blind persons are not able to see the meaningful movements, and facial gestures of the participants. The formal meeting minutes and / or participants' conversation during the meeting normally lack this important feedback in order to determine who is in favor and who is against their proposed suggestions. This is crucial in business negotiations, where one has to convince people and do lobbying for winning the business case in upcoming meetings. Today the devices already exist for instantly and seamlessly capturing the snapshots everywhere. The proposition suggests data capture using a similar device called SenseCam[1], and then making these snapshots accessible for the benefit of the visually impaired users.

1 Introduction

The archiving of meeting proceedings is considered to be very useful for the immediate benefit of the participants. However, exploring the associations between the meeting room constituents (participants, & objects within the meeting room) vastly increases its benefits. Multiple forums & vocabularies, & Multi-media information integration are marked as two of the technical research challenges by [1]. The important clues for meeting recognition are *What is being discussed, Who is in action, To whom one is talking to, When the meeting is taking place, & Where* [2].

The above questions can be significantly answered by managing the associations of user's information items using ontologies. Our prototype personal information management system SemanticLIFE [3] explores & manages the various types of associations existing within user's lifetime information using semantic web technology. *Firstly, there exist structural associations*, i.e., each information item has an inherent association with its structural metadata. For example, an email is associated with header fields such as Subject, From, To, Received / Sent Date, etc. *Secondly, associations could be asserted using manual annotations.* For example, a contact information item x can be manually associated with a project proposal document y based upon collaboration done by x on y. *Thirdly, using sophisticated techniques of textual and multimedia content analyses* further associations in the concerned ontology are

[1] http://research.microsoft.com/sendev/project_sensecam.aspx (5th Mar'06).

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 1147 – 1154, 2006.

possible. The examples are the possible association of an email message containing the word “ICCHP” in body with some web page having title “ICCHP 2006”, association of information items existing in the same time slice (editing a project proposal document & browsing the web for related information in parallel), association of information items generated from the same location or related with same location (a picture taken & a telephone call made from the same place). Also, significant benefit is achievable by finding the active / passive project participants based upon the presence / absence in related project meetings which can be calculated once the pictures are annotated for relative movements / gestures of the participants.

The next step is to present these associations in an accessible way for people with special needs. Generally for this purpose, accessibility is taken into account at the presentation layer according to the Web Content Accessibility Guidelines (see WAI[2]). The guidelines as their name suggests are primarily focused on the content of web pages which is mostly textual HTML based information. The scenario in the case of our prototype SemanticLIFE is different though. Because the data feed is not only limited to web pages. The primary contents in our system are the user information items such as emails, documents of various types, chat sessions, telephone logs, process state data, browsed web history, contacts, calendar data & snapshots to be captured by SenseCam or some other similar device. When the associations are visualized at the interface then the appropriate accessibility guidelines can be applied.

The point of our interest is to apply the accessibility criteria also at the contents (making accessible the associations of captured snapshots) & not just at the presentation level for facilitating the exploration of associations for people with special needs.

In the next section related work is discussed. Then some sample scenarios are described which we intend to fulfill, followed by proposed approach and workflow for managing associations.

2 Related Works

The importance of the domain is highlighted by the ongoing projects related with meeting room recognition technologies such as Meeting Recorder Project by [4], Automatic Meeting Recognition Project by [1], Meeting Browser and Computers in the Human Interaction Loop by [5], Augmented Multiparty Interaction [6], and Interactive Multimodal Information Management [7], just to name a few. Multimedia techniques are used in [2] to track the meeting with participant’s ID using color appearance, face id, and speaker id. The corpus based framework in [8] describes how the meetings are modeled in layers, and how the annotation could be used for meeting recognition. The work done by [9] is highly related with our work. They have used image processing techniques for monitoring and tracking the user activities. To achieve the same goal we are focusing on linking the user’s information items with each other based upon metadata. This will also help in resolving the synchronization issue between different data sources, a problem mentioned by [9]. The dialog act labeling guide [10] describes the audio dialog structure & the annotation system used.

---

[2] Web Accessibility Initiative of the W3C `http://www.w3.org/WAI/` (5^{th} Mar’06).

Our approach is different since it is concerned with usage of annotations instead of audio dialog for exploiting the gestures and participants' movements. It is possible to identify and annotate the meeting room constituents by using ontologies of interconnected information items. We are optimistic that this component will supplement the capture and recognition of meeting room knowledge exchange significantly.

Examples of available image annotation tools for the semantic web are flick2rdf[3], PhotoStuff[4], SWAD[5], M-OntoMat-Annotizer[6]. Some of these convert the annotations into owl file like flick2rdf. The supported annotations are mostly user comments and are not sufficient for our purpose of capturing the metadata. The recent use case document by W3C [11] describes the issues and challenges in carrying out manual, semi-automatic, and completely automatic annotation of images. Our approach is in-line with their vision as we are starting with semi-automatic annotations and keeping our system architecture flexible to integrate the efforts done by image processing community in future.

For applying our strategy the meeting snapshots must be captured in sufficient detail. SenseCam which is a badge size wearable camera is capable of doing this under user control. It can take pictures of the meeting room based upon any small change in environment (location of objects, light & temperature). It is in use by Microsoft in their research project MylifeBits [12] with promising results highlighting new challenges in managing personal information in various domains. *We intend to use it for blind by exploiting the annotation mechanism in our prototype SemanticLIFE.*

3 Sample Scenarios

Blind person sitting in a business meeting: The possible movements or postures by the participants can be leaving or entering the room, sitting down, standing up, whispering with someone while leaning, relaxing on the chair, sitting alert, hand gestures by the participants, talking on telephone, working on laptop, apparently sleeping etc. These are some of the meaningful movements along with facial gestures to estimate the mood of the participants during discussion.

- The blind person wears the SenseCam during the meeting. So the pictures of the whole proceeding are taken automatically.
- Pictures are uploaded in SemanticLIFE repository as our file upload data feed.
- Retrieval of day's pictures from the repository and identification of participants either manually by the caregiver or automatically [13] using multimedia analysis plugins with possible help from ontologies.

3 `http://purl.org/net/kanzaki/flickr2rdf` (5th Mar'06).

4 `http://www.mindswap.org/2003/PhotoStuff/` (5th Mar'06).

5 `http://swordfish.rdfweb.org/discovery/2004/03/w3photo/annotate.html` (5th Mar'06).

6 `http://www.acemedia.org/aceMedia/results/software/m-ontomat-annotizer.html` (5th Mar'06).

- Annotation of pictures by the caregiver based upon the gestures / movements of the constituents (participants & other objects in the vicinity). Initially, it is assumed that the caregiver is already informed about the identities of the constituting objects. Later, the identities should be matched against similarity using ontologies.
- Enrichment of associations; For example, the gestures by a participant would update his / her contact profile for a particular project meeting. This will give useful information to the blind user about this participant in future meetings.

The system is usable in domains other than project meetings & also for a more diverse range of special needs [14]. The following scenario is useful for blind persons as well as for mobility impaired users. However, this is not the main focus for this paper.

Blind person visiting the city: In this case, the captured data is much broader. The pictures of our interest can also be those of stationary objects like restaurants, information counters, monuments, facilities coming on way to the destination.

- Blind person equipped with SenseCam is traveling from one station to another. The pictures taken would most probably include the facilities like restrooms, lifts, ramps about which the blind person is not aware of on his / her initial visit.
- Similar to the above scenario, the pictures would be annotated by his / her caregiver, & associations would be made using ontology.
- Before making the subsequent visit, the blind user can consult his / her route planning. Based upon the associations made in the previous step, the system could present to him the information about availability of accessible facilities en route and now the travel planning & travel itself could become more accessible.

Since, the SenseCam capture rate is about 1 picture/sec. Therefore, for an hour proceeding, there are 3600 pictures which is a big number. The annotation mechanism should be very user friendly so that the caregiver easily annotates all the pictures.

4 Proposed Approach

The big amount of pictures is categorized into distinct parts based upon meaningful criteria such as given in Fig.1. The parts thus holding a smaller amount of photos are annotated which is a relatively manageable task. The criteria to categorize the parts are devised in such a way so as not to loose any meaningful information about the whole meeting proceedings, yet being able to retrieve everything based upon a group of criteria, like a data cube [15]. The primary criterion is time distribution. The secondary criterion is to change the part boundaries based upon the participants' tracking (their identification and relative postures). Another criterion is to associate the issues or agenda items discussed during the meeting with the parts. There can be more criteria in future, based upon evolving semantics & user needs.

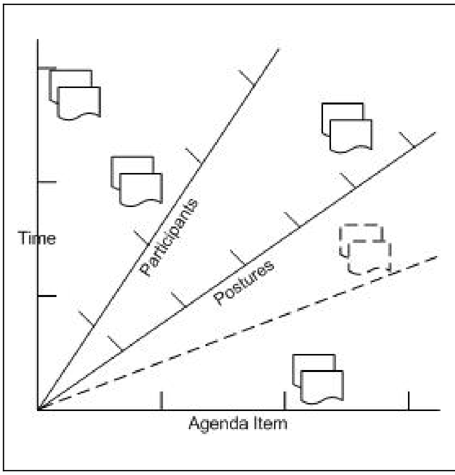


Fig. 1. Information distribution over multiple axes

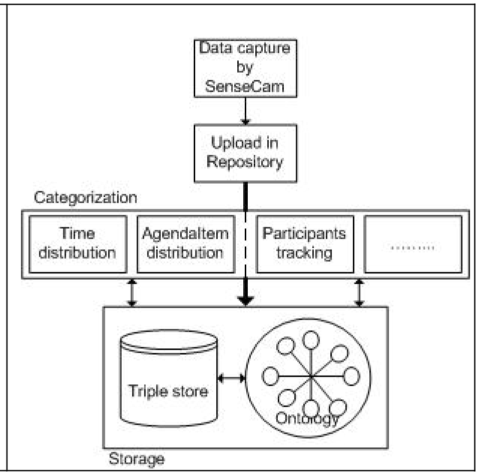


Fig. 2. Workflow of annotation subsystem

Already existing project related ontologies (e.g., UMBC[7], doap[8]) are unable to fulfill our requirements, but can be extended. *We propose an ontology (see Fig. 3) for specifying concepts related with projects, meetings & the binding between them.*

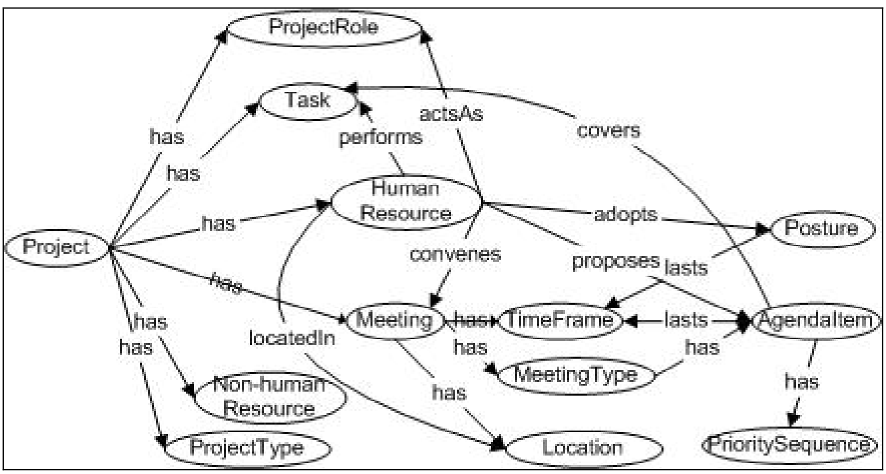


Fig. 3. A sample Project Meeting Ontology

The ontology is in its initial state. However it depicts the required concepts to fulfill our scenario. "MeetingTypes" can be highly formal, moderately formal, informal,

[7] `http://ebiquity.umbc.edu/ontology/project.owl` (5th Mar'06).
[8] `http://usefulinc.com/doap` (5th Mar'06).

research brainstorming, regular departmental meeting, coordination meeting etc. Meetings are carried out in context of some project(s). “ProjectType” can be software development, construction, tourism, entertainment etc. A “ProjectType” has specific “Roles” such as manager, analyst, developer, tester, and integrator. depending upon the project responsibilities. “Tasks” are associated with “Roles” to carry out specific responsibilities, and are discussed in meetings under specific “AgendaItems” lasting some “TimeFrame”. The meeting agenda is typical for specific type of project and specific type of meeting. Meeting is convened by someone, & the agenda items are also proposed by someone. The meeting has a certain “Location”. The person also has a location while attending the meeting (inside the meeting room or at remote location), the meeting has a “TimeFrame”, & so do the agenda items which are normally prioritized (have “PrioritySequence”). Each “Agenda Item” lasts for some duration, & the participants assume “Postures” lasting some timeframe throughout the meeting.

Distribution according to time: A long meeting can be heuristically broken down into n minute’s duration of m parts each. For this purpose the timestamp available in Exif header of the pictures is used. The process is carried out automatically by the analysis plugin in our system based upon predefined values in the configuration file.

Distribution according to participants’ tracking: The tracking of participants is very important because the issues discussed / decided in the meeting can be meaningfully related with it. There can be two steps, i.e., person’s identification and posture‘s identification. For example, distribution of pictures based upon names of the persons. Initially it is sufficient to tag the exit / entry of a participant which is possible by following a manual protocol, or using RFIDs. Later, some of the major postures lasting for the duration of parts can be identified & annotated.

Distribution according to agenda items: The nature of the AgendaItems is specific from project to project. Some agenda items will be common in many projects. There is a small chance that the items are unique for a common project & meeting type.

5 Workflow for Managing Associations

The three types of associations (structural, manual, dynamic) mentioned in Section 1, are taken care of for the stored pictures in order to make these accessible for blind. The information items are stored as RDF triples in SemanticLIFE repository against a base ontology. This ontology specifies the header fields for each information item as its properties. The UI plays an important role for convenient interaction & annotation of pictures. The different types of information items from this triple store are displayed differently, say in different colors & on different graph. According to user studies the whole day is divided into 4 to 15 activities out of which meetings are generally 2 to 5 for users under study. The activities with their times & durations may already be described in the calendar. Using the range sliders user is able to select slices of timeline during which a meeting has taken place. At times, the planned activities overlap with performed activities. Then the user can handle it by adjusting & confirming the slider positions on the UI. In the selected time slice user can filter out the items other than the pictures. Each selected time slice may correspond to Subject of the appointment in calendar. If not specified in calendar, then a right click of the mouse would enable the user to put the items of the selected time slice into a new named collection, say SlifeMeeting of type “Meeting” which may be connected to Meeting ontology specified elsewhere on the web. Participants for this meeting are retrieved from calendar, if specified already. Otherwise, right clicking the SlifeMeeting, would enable the user to enter meeting participants.

Structural enhancement of information items: By visualizing the retrieved pictures components of the meeting room constituents are described by the caregiver. The picture Exif header specifies Camera-specific properties (make, model, sensing method, lens size etc.) & data about Image-specific properties (creation date, image resolution, height, width etc.). The structure can be enriched to take into account the constituents, e.g., describe who is present in the picture, & their postures. This is laborious & time consuming task. By using intuitive UI, & ontology of interconnected information items, the task is made more suggestive & convenient for the user.

Manual associations with other information items: It is possible that some other activity like a telephonic conversation, chat session or web browsing was carried out during the meeting. These activities may be related with the meeting under progress. Using manual annotations the individual pictures or collections of pictures can be associated with each other or any other information item in the repository or a concept,e.g., rating some named collections as highly / moderately useful or useless. It will be beneficial for analyzing the time usage by you & other meeting participants.

Dynamic associations: Once the ontology (see Fig 3) instances are populated, then dynamic association of entities is possible. The information extraction techniques will be applied to the contents of fed information in cases where linking of fields in item header is insufficient to firmly establish the associations. For this purpose concepts / key-terms are mapped with the ontologies. The links established through the ontology would make it possible to see the historical behavior of a person in specific type of meetings, analysis of specific agenda items with the duration of discussion etc.

6 Conclusions and Future Work

The Automatic data capture devices like SenseCam do have the capability to capture much of user’s activities. But, it is still far from building an automatic diary for the user due to missing associations. An ontology of information items is developed for our prototype. A proposal is suggested to capture the meeting snapshots & make those accessible for blind. A projects meeting ontology is proposed which would be of particular benefit for blind people & generally beneficial for all. The work is in progress. Many privacy issues are to be investigated once the initial corpus is built.

Acknowledgements

This work is supported by the financial support of Austrian National Bank (OeNB) for Business Information Management research project, the ASEAN-EU University Network Initiative, & the Higher Education Commission of Pakistan.

References

1. NIST Automatic Meeting Room Recognition Project, http://www.nist.gov/speech/test_beds/mr_proj/index.html (5th Mar'06).
2. Yang J., Zhu X., Gross R., Kominek J., Pan Y. and Waibel A., Multimodal People ID for a Multimedia Meeting Browse, Proceedings of the ACM Multimedia 1999, pp. 159-168.
3. Ahmed M., Hoang H.H., Karim M.S., Khusro S., Lanzenberger M., Latif K., Michlmayr E., Mustofa K., Nguyen H.T, Rauber A., Schatten A., Tho M.N. and Tjoa A.M., 'SemanticLIFE' - A Framework for Managing Information of A Human Lifetime, Proceedings of the International Conference on Information Integration, Web-Applications and Services, IIWAS'04, 27-29th Sep. 2004 Jakarta-Indonesia.
4. ICSI Meeting Recorder Project, http://www.icsi.berkeley.edu/Speech/mr/ (5th Mar'06).
5. interACT Meeting Room Project, http://penance.is.cs.cmu.edu/meeting_room/(5th Mar'06).
6. Augmented Multi-Party Interaction, http://www.amiproject.org/ (5th Mar'06).
7. NCCR Interactive Multimodal Information Management, http://www.im2.ch/ (5th Mar'06).
8. Reidsma D., Rienks R.J. and Jovanovic N., Meeting modeling in the context of Multimodal research, Proceedings of the MLMI'04, LNCS 3361 (2005), pp. 22-35.
9. Wactlar H., Bharucha A., Stevens S., Hauptmann A. and Christel M., A System of Video Information Capture, Indexing and Retrieval for Interpreting Human Activity, IEEE International Symposium on Image and Signal Processing and Analysis (ISPA'03), Special Session on System Perspectives in Information Retrieval, Rome, Italy, September 18-20, 2003.
10. Elizabeth S., Raj D., Sonali B., Jeremy A. and Hannah C., The ICSI Meeting Recorder Dialog Act (MRDA) Corpus, Proceedings of SIGDIAL'04 (5th SIGdial Workshop on Discourse and Dialogue), 2004, pp. 97–100.
11. Image Annotation on the Semantic Web, http://www.w3.org/2001/sw/BestPractices/MM/image_annotation.html#use_cases
12. Gemmell J., Bell G. and Lueder R., MyLifeBits: a personal database for everything, vol. 49, No. 1 Communications of the ACM, January 2006, pp. 88-95.
13. Jeon J., Lavrenko V. and Manmatha R., Automatic image annotation and retrieval using cross-media relevance models, Proceedings of the 26th International ACM SIGIR Conference SIGIR 2003, ACM 2003, pp. 119-126.
14. Pühretmair F., It's Time to Make eTourism Accessible, Proceedings of the 9th International Conference on Computers Helping People with Special Needs ICCHP 2004, LNCS 3118 (2004), pp. 272–279.
15. Gray J., Chaudhuri S., Bosworth A., Layman A., Reichart D., Venkatrao M., Pellow F. and Pirahesh H., Data Cube: A Relational Aggregation Operator Generalizing Group-by, Cross-Tab and Sub Totals, Data Mining and Knowledge Discovery Journal, vol. 1, No. 1, 1997, pp. 29-53.

On the Accuracy of Tactile Displays

Christopher Power

Department of Computer Science
The University of Western Ontario
London, Ontario
N6A 5B7, Canada
power1@csd.uwo.ca

Abstract. Inaccuracies in the reporting of finger positions from electronic tactile displays can result in errors in the audio presentation of multi-modal applications. In this paper, we conduct an experiment to examine the accuracy of one such device. Given the results of this experiment, we present a collection of recommendations for the spacing of objects within a tactile scene.

1 Introduction

We consider a set of multi-modal applications using finger tracking technology on a tactile display to deliver timely, relevant audio information regarding tactile features encountered by a blind user. Such applications rely on finger tracking technology such as a touch pad [1,2] or a sensor device in which the finger is placed [3]. In either case, the application must identify the tactile features detected by the fingers of a user and associate those features with the speech and non-speech sounds.

In this paper, we examine how prevalent inaccuracies in positional readings are in the exploration tasks of users using such devices. We conduct an experiment which simulates the types pointing and steering motions in which a user partakes while she/he is exploring a tactile scene. During this experiment, we compare the tactile feature for which the user is searching to the positional information reported by the device.

Using the data from these experiments, we provide recommendations regarding the constraints for presenting tactile graphics on such devices.

2 Related Research

There are a variety of devices which have been, or are now, available to the visually impaired user group for the presentation of tactile graphics.

The most well known, and some say the most advanced [4], tactile presentation device was the Optacon [5,6]. This device scanned the surface of various objects, such as coins or grocery receipts, to create a two dimensional tactile representation. This device was well received by the populace, but unfortunately it was discontinued in 1997.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 1155–1162, 2006.

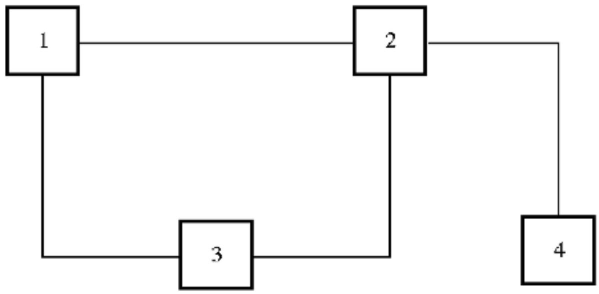


Fig. 1. A simple graph

Recently, touch pad devices, such as the Nomad Board used in the Audio-Touch project [7], the Talking Tactile Tablet (TTT) from TouchGraphics [1], and the IVEO from ViewPlus [2], provide the ability to add audio annotations to static tactile documents. With these devices a user places a static overlay, produced through any of embossing, theroform or microcapsule technolgies, on the touch pad of the device. When the user touches the pad, the finger location is transmitted to a personal computer which plays associated speech and non-speech sounds for the user.

This multi-modal approach of presenting information is also possible through refreshing pin displays. These devices create raised pictures by raising a two dimensional array of pins. Examples of these devices include the prototype Metec Dot Matrix Display (DMD) [3] or the dynamic pin display from NIST [8]. Such devices track finger positions via finger sensors which report the region, or point, of the pin board over which the sensor is placed.

3 Problem Description

Now, we discuss the impact that inaccurate reporting can have on the exploration of a multi-modal document. For purposes of this discussion we refer to the the simple undirected graph presented in Figure 1. Assume that the nodes are labeled with audio tags consisting of the number of the node[1]. As such, when a multi-modal application detects the user's finger within the boundary of the node, the label will be spoken through a speech synthesizer.

Assuming that the document is presented on our sample device, the Metec DMD, the user's finger position is detected via a sensor ring in which the finger rests. This finger sensor has a diameter of 8 dots, and we assume the sensor returns the center of the ring as the finger position.

The Figure 2a depicts the ideal situation: the user's finger, represented by a cross, is in the center of the ring. In this situation, the multi-modal application would correctly interpret the user's target and speak the number associated with

[1] The numbers listed in the diagram are for the reader's benefit and may not be present on the tactile diagram.

it. In the case of Figure 2b we have the user's finger within the, but the center point of the ring is outside of the boundary. Given this situation, there would be no audio feedback from the multi-modal application despite the user having are correctly pointed at a target.

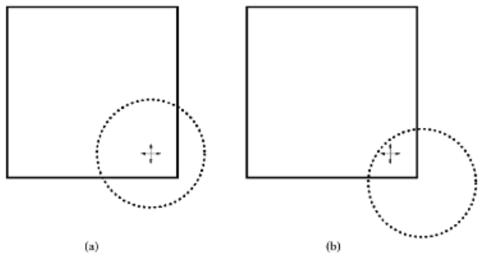


Fig. 2. (a) The sensor ring centered over the finger position. (b) The sensor ring offset from the finger position

Similar accuracy problems have been observed in touch pad devices. An empirical evaluation of the TTT [1] device showed that it reported accurate information on self-authored graph diagrams approximately 80% of the time. It is hypothesized that such inaccuracies come from variations in the use of the full area of third metacarpal row of the finger to apply pressure, versus using only the fingertip. Such inaccuracies were particularly prevalent when dealing with narrow targets such as line segments.

As these types of inaccuracies are found in many of the current technology options available for multi-modal presentation, we explore how prevalent such inaccuracies are likely to be during tactile exploration tasks.

4 Methodology

We propose a series of experiments intended to determine how often errors occur in reporting targets pointed at by the user.

4.1 Subjects

There were 14 volunteer subjects for our experimental trials, seven male, and seven female. The subjects were between 20 and 60 years of age. All subjects were right handed and used their dominant hand during the test. Each subject was fully sighted; however, due to the nature of the DMD, she/he was unable to see the picture, and thus she/he had to rely on her/his sense of touch for identifying features.

4.2 Apparatus

For the tactile image, we used the Metec Dot Matrix Display (DMD) 120060, which contains a surface of pins, 120 pins wide and 60 pins in height. These pins can be moved independently from one another, resulting in a raised picture which can be explored with a subject's hands. Each pin is approximately the size of a Braille point, which results in a low resolution image of approximately 10 points per inch.

Attached to the DMD are two finger sensors, one for each hand. For this experiment, we only use one of these sensors, with the subject's index finger placed within the detection ring of the sensor. Contained within this ring is a electrical coil which detects the position of the sensor on the pin board. In this experiment we sample all pins contained within the sensor ring and use the center of gravity for the sensor as the position of the user's finger.

We recorded the finger positions of the user on a Windows based computer with a Pentium 4 1.2 GHz processor with 512 MB of RAM. The application for user interaction was written in the Java 5 programming language.

4.3 Procedure

A variety of diagram styles were used to simulate the *pointing* [9] and *steering* [10] tasks of a user exploring a tactile graphic.

The first diagram type is depicted in Figure 3. In this diagram a subject started her/his trials with her/his index finger on the center rectangle. Each subject alternated between pointing at a target in the outer ring and traveling from one of those outlying targets back to the center square.

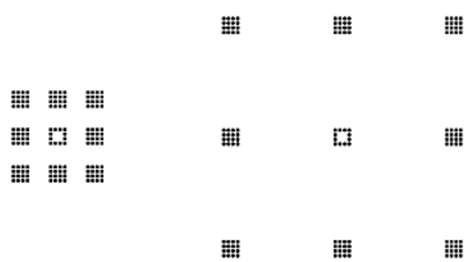


Fig. 3. Pointing Based Experimental Diagrams

Each target was four pins by four pins, resulting in an object which was as tall as a eight-point Braille character. The diagrams differed in the distance of the outer ring from the center square, with the distances of the squares at the cardinal points of the five diagrams being 4, 8, 12, 16 and 20 pins.

For the steering tasks we used two diagrams. The first was the simplest of steering tasks: the straight line. On this diagram the subjects started at the left hand side of the line and alternated moving to the right and to the left, steering along the line, and stopping at the ends. Each line was 100 pins in length, with line widths of 1, 2 and 4 pins. Stopping at a position on the line within ten dots from the end was considered a successful steering trial.

In order to have the subjects perform a more complex steering task, a set of step diagrams were included in the test. The steps had the widths of 1, 2 and 3 pins and were tested with distances traveled of 48, 63 and 101 pins respectively (measuring the topmost edge of the steps). The trials alternated from the top of the steps to the bottom and vice versa. The top or bottom horizontal bar was the target for the user to find depending on the trial. For all steering tests the users were asked to follow the line as close as possible.

When a diagram was first presented to a subject, she/he was given the opportunity to familiarize herself/himself with the tactile picture through finger exploration without the finger sensor. This familiarization phase was intended moderate training effects that may come from a subject becoming more familiar with a particular diagram as the test progressed.

The users were instructed to place their index finger in the sensor ring. For each trial on a diagram the users were instructed to:

1. Listen for a tone to signal the beginning of a trial.
2. Move the finger to the target listed on the screen.
3. When the designated target is arrived at, stop movement over the target.

There were 20 trials on each diagram to further reduce the effects that training may have had on the recorded data.

If the subjects asked "Should I be concerned with time or accuracy?", the experiment supervisor was instructed to respond "That is your choice."

5 Results

Obvious target errors, such as pointing at the wrong end of the horizontal line, were removed by inspection. Thus, in the results, we assume that after data cleaning all data points indicate a trial where the subject has successfully placed her/her finger on the intended target. Given this fact, we examine whether or not the position reported by the sensor reflects this fact.

The Table 1 show the confidence intervals for the proportion of trials in which the point recorded was within the boundary of the intended target, which is referred to in the sequel as a "hit".

For the pointing task, we see that the success rate of identifying the target via the position of the sensor is quite high, with a sample proportion of approximately 88% and a narrow 95% confidence interval. Using this result, we can interpret the result as follows: the predictions in our sample were successful approximately 88% of the time. Further, we are 95% confident that the true unknown proportion of successful predictions lies somewhere between 86 to 90%.

However, when we examine the steering task results, we find that there is a very low proportion of the trials where the reported position of sensor was within the boundary of the target. Additionally, there was negligible improvement as the paths became wider. The result in the last row of Table 1, where the target is equal in height to the pointing task targets, is particularly surprising. We would expect to see a sample proportion similar to the pointing task for this object.

Table 1. Confidence Intervals for Steering Trials

Interaction Type	Target Height	Hits	Trials	Sample Proportion	95% Confidence Interval
Point	4	1234	1400	88.1	86.4–89.8
Steer	1	30	515	5.8	3.8–7.8
Steer	2	83	516	16.1	12.9–19.3
Steer	3	99	260	38.1	32.2–44.0
Steer	4	105	280	37.5	31.8–43.2

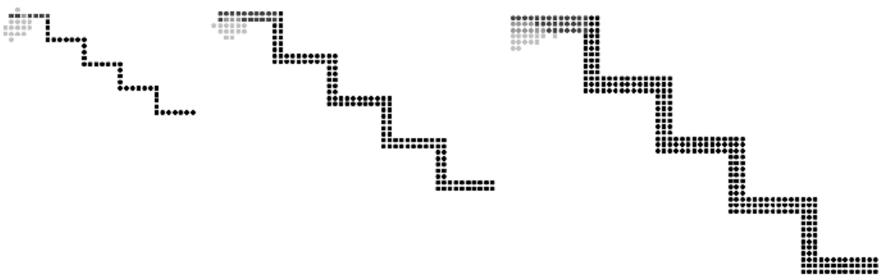


Fig. 4. A visualization of the steering results on the step shaped diagrams

Table 2. Sample Proportion and 95% Confidence Intervals for Subject Trials with Flexibility

Interaction Type	Target Height	Flexibility (η)	Hits	Trials	Sample Proportion	95% Confidence Interval
Point	4	1	1369	1400	97.8	97.0–98.6
Point	4	2	1385	1400	98.9	98.4–99.5
Steer	1	1	257	515	49.9	45.6–54.2
Steer	1	2	476	515	92.4	90.1–94.7
Steer	2	1	325	516	62.9	58.8–67.2
Steer	2	2	491	516	95.1	93.3–97.0
Steer	3	1	193	260	74.2	68.9–79.5
Steer	3	2	247	260	95.0	92.4–97.6
Steer	4	1	221	280	78.9	74.2–83.7
Steer	4	2	272	280	97.1	95.2–99.1

As this is not the case, it leads us to the question: what is different about the steering task which results in such a lower proportion?

When we examined the steering trials, we found that the points recorded by the device fell consistently off an edge of the steering path. A visualization of this phenomenon on the step diagrams is depicted in Figure 4. This pattern lends evidence to the proposition that the majority of the users use an edge of the path, as opposed to the center of the path, for steering purposes.

Due to this pattern, we choose to expand our definition of what is considered to be a hit in a trial. For any recorded value from the sensor, we will consider it to be a hit if it is within η dots of the edge of the tactile object. We refer the value η as the *flexibility* in our definition of what is considered on target.

Table 2 shows the same data as above evaluated with differing levels of flexibility.

There is a significant improvement in the accuracy of the readings in terms of points being reported within a target. The results at the flexibility level of 2 have sample proportions in the range of 92 to 97% for all line widths, all with narrow 95% confidence intervals.

6 Discussion

These results indicate that a multi-modal application which is intended for use on a device such as this requires some flexibility in how it interprets results from the sensing mechanism. In order to compensate for this flexibility, we can make the following recommendations regarding the preparation of tactile scenes:

1. Objects that are of size 4 pins by 4 pins or larger are sufficient for accurate reporting of target information for the majority of pointing tasks.
2. A distance of 5 pins between objects is sufficient to distinguish between objects in the majority of pointing tasks, even with two pins of flexibility.
3. Accurate reporting of targets in steering tasks does not substantially improve as lines get wider. This implies that thin lines can be used somewhat liberally in tactile diagrams.
4. Parallel lines which serve as paths within a tactile diagram must be spaced at least 5 pins distant to avoid confusion in audio information presentation.

7 Conclusion

In this paper we have presented an experiment examining the accuracy of a refreshing pin display that can be used to present a multi-modal document to a visually impaired user.

This experiment showed that the finger position reported by the display did not always correspond to the finger position as perceived by the user. Such correspondence was highest during pointing tasks with the device reporting a position within the target at which a user was pointing approximately 88% of the time. However, steering tasks fared far worse with extremely low rates in target correspondence, even when the lines were quite thick. When we expanded the definition of what was on target to include points outside of the intended object, it was found that a flexibility of two pins was enough to get a high proportion of correspondence between intended targets and position reporting.

Using these results, we have provided a collection of recommendations which are intended as lower bounds for spacing of objects in a tactile scene designed for presentation on such interactive devices.

References

1. Landau, S., Gourgey, K.: Development of a Talking Tactile Tablet. Information Technology and Disabilities **7**(2) (2001) Online Publication Retrieved from `http://www.rit.edu/~easi/itd/itdv07.htm` (09/2005).
2. ViewPlus: IVEO. Online Product Web Site Retrieved from `http://www.viewplustech.com/` (09/2002) (2005)
3. METEC: Braille-großdisplay dmd 120060. Ingenieur-Gesellschaft mbH, Stuttgart (1989)
4. Stein, B.K.: The Optacon: Past, Present, and Future. DIGIT-EYES:The Computer Users' Network News (1998)
5. Juillard, J.: Optacon helps the blind to read. Schweizerische Technische Zeitschrift **79**(2) (1982/01/27) 18 – 19
6. Ikei, Y., Wakamatsu, K., Fukuda, S.: Vibratory tactile display of image-based textures. IEEE Computer Graphics and Applications **17**(6) (1997) 53–61
7. Lötzsch, J.: Computer-aided access to tactile graphics for the blind. In Zagler, W.L., Busby, G., Wagner, R.R., eds.: Computers for handicapped persons : 4th International Conference, ICCHP '94, Vienna, Austria, September 14-16, 1994. Volume 860 of Lecture Notes in Computer Science., New York, NY, Springer-Verlag (1994)
8. Roberts, J.: NIST Refreshable Tactile Graphic Display: A New Low-Cost Technology. In: Proceedings of the 2004 CSUN Conference on Technology and Persons with Disabilities, Los Angeles, CA, California State University Northridge (2004) Online Publication Retrieved from `http://www.csun.edu/cod/conf/2004/proceedings/csun04.htm` (09/2005).
9. Fitts, P.: The information capacity of the human motor system in controlling the amplitude of movement. Journal of Experimental Psychology **47** (1954) 381–391
10. Accot, J., Zhai, S.: Beyond Fitts' Law: Models for Trajectory-Based HCI Tasks. In: Proceedings of ACM CHI 97. (1997) 295–301

HOWARD: High-Order Wavefront Aberration Regularized Deconvolution for Enhancing Graphic Displays for Visually Impaired Computer Users

Miguel Alonso Jr.[1], Armando Barreto[1], Malek Adjouadi[1], and Julie A. Jacko[2]

[1] Florida International University, 10555 West Flagler Street, Miami, FL 33184, USA
{malons05, barretoa, adjouadi}@fiu.edu
http://dsplab.eng.fiu.edu/DSP/index.html
[2] Georgia Institute of Technology, 765 Ferst Drive, NW, Atlanta, GA 30332, USA
julie.jacko@isye.gatech.edu
http://www.isye.gatech.edu

Abstract. High-Order Wavefront Aberration Regularized Deconvolution (HOWARD) is a complete closed loop system developed for simulating human visual function with the primary goal of enhancing graphic computer displays for users that have refractive errors (resulting in difficulty interacting with visual displays). Visual function is a primary requirement for a human being to engage in computer usage efficiently. There are situations in which common forms of vision correction, such as contact lenses or glasses are not sufficient to provide the necessary compensation for some users to interact with graphic displays. This paper presents a model for the visual function of an imaging system, the implementation of an artificial eye with high-order wavefront aberrations, as well as a method for providing compensation of the artificial eye through a graphic display.

1 Introduction

Since the development of the Hartmann-Shack sensor for measurement of wavefront aberrations in the human eye [5],[7], it has become possible to characterize the individual refractory limitations of each individual. This capability has been used by emerging techniques such as Adaptive Optics [8] to implement custom vision correction approaches. These methods, however, may be expensive and cumbersome due to the highly specialized hardware needed. The corrective approach we propose requires only the knowledge of the wavefront aberration to create customized pre-compensated images to be displayed in standard PC graphical displays to facilitate the interaction of users with refractory limitations.

1.1 Problem Statement

The goal of this paper is to present an innovative way of providing enhancement of graphic displays to PC users that have high-order visual aberrations. A number of

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 1163 – 1170, 2006.

visual impairments, such as keratoconus [10], involve high-order visual aberrations that may not be corrected by using glasses or contact lenses.

The algorithm presented here relies on the linear systems approach to modeling the human visual system, known as Fourier Optics [6]. The human visual system can be thought of as a linear system having an impulse response *H*. In a linear system, the output of the system is the convolution of the input with the impulse response of the system. The impulse response of an ideal optical system, including the human eye, is a delta function. Thus, if the user is free from any visual aberrations, his/her eye impulse response, from here on termed Point Spread Function (PSF), will be a delta, allowing the user to interact more efficiently with the personal computer (PC) via the graphical display. This will result in a clear, undistorted projection of the object onto the retina. If however, the user has a visual aberration, the PSF will not be a delta, and thus the retinal project of the object will be distorted. Fig. 1 shows the Linear Shift Invariant (LSI) model used to describe the optical process.

An object $O(x,y)$ (for example, a picture on a graphical display) is degraded by convolution with the PSF of the user's visual system, $H(x,y)$, resulting in a distorted projection of the object on the user's retina, $I(x,y)$. This is described by

$$I(x,y) = H(x,y) * O(x,y), \quad x = 0,\ldots,N-1, y = 0,\ldots.M-1 \quad . \tag{1}$$

where * denotes convolution.

Given *O* and *H*, the High-Order Wavefront Aberration Regularized Deconvolution (HOWARD) algorithm seeks to find an inverse function H^{-1} to produce an enhanced object, *EO*, counteracting the distortion introduced by *H,* such that when the user views the *EO* on the graphic display, an undistorted version of *O* will be projected onto the retina.

$$EO(x,y) = H^{-1}(x,y) * O(x,y) \quad . \tag{2}$$

$$I(x,y) = H(x,y) * EO(x,y) = H * H^{-1} * O(x,y) \approx O(x,y) \quad . \tag{3}$$

This model amounts to a noiseless deconvolution problem. However the ill-conditioned nature of the PSF of the human eye [1] will require a robust method, such as HOWARD, to allow a more efficient interaction between a user that has a refractive error present in his/her visual system and the PC.

2 Background on Fourier Optics

In order to better understand HOWARD, a brief review of Fourier Optics, the foundation for why the human eye can be considered a Linear Shift Invariant System, will be given.

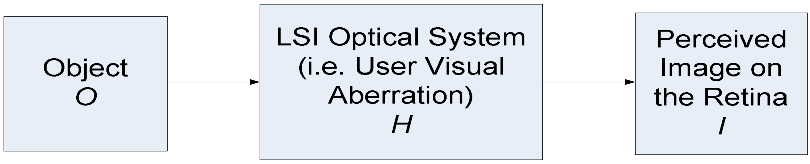


Fig. 1. Simplified Human Visual Model. The image perceived by the user results from the convolution of an object, in this case an image on a graphical display, with the Point Spread Function (PSF) of the user. Under ideal conditions, the perceived image on the retina will be a magnified, but undistorted version of the object. If the user has any type of visual aberration, the resulting image that falls on the retina will be a distorted version of the object.

The foundation of Fourier Optics lies in the definition of the irradiance pattern that falls at the focal point of an optical system composed of a lens and an aperture:

$$E(X,Y) = \iint_{aperture} A(u,v) e^{ik(Xu+Yv)/F} dxdy = H, \quad . \tag{4}$$

where *E(X,Y)* is the irradiance pattern (which is, in fact, the point spread function of the system [12],[13]), *A(u,v)* is the pupil function representing how a plane wavefront from a source at infinity exits the pupil towards the imaging plane, and F is the focal distance of the lens [13]. In equation (4) we can see that this is in fact a Fourier transform. The PSF of any optical system is the Fourier transform of the pupil function. We refer the reader to [6], which details the relationship needed to apply linear systems theory to optical systems.

Most optical systems, including the human eye, have a complex aperture that modulates both the intensity and the phase of the incoming plane wavefront. For an ideal optical system, the pupil function will be one (*A(u,v)* = 1) across the entire pupil, causing no modulation of intensity or phase. A user that has refractive errors will not have an ideal pupil function. Instead, the pupil function will be composed of a real and imaginary part, as follows:

$$A(u,v) = D(u,v) e^{-i2\pi n W(u,v)/\lambda}. \quad . \tag{5}$$

where *D(u,v)* represents the intensity aberration function and *W(u,v)* represents the wavefront aberration function, n is the index of refraction, and λ is wavelength of the light that is incident on the optical system. For a human eye that only has refractive aberrations, i.e., does not have a disease such as cataracts that disperses the incoming light, $D(u,v) \approx 1$ [12].

The wavefront aberration function provides knowledge of how an optical system distorts the incoming wavefront. In order to proceed with any type of enhancement, knowledge of the wavefront aberration function is necessary on a custom basis for each user. Developments in ophthalmology during the last decade have made access to a user's wavefront aberration in the form of Zernike polynomials possible [14], [7]. Once this is known for the user's eye, the PSF describing the way the user views the graphical display can be calculated.

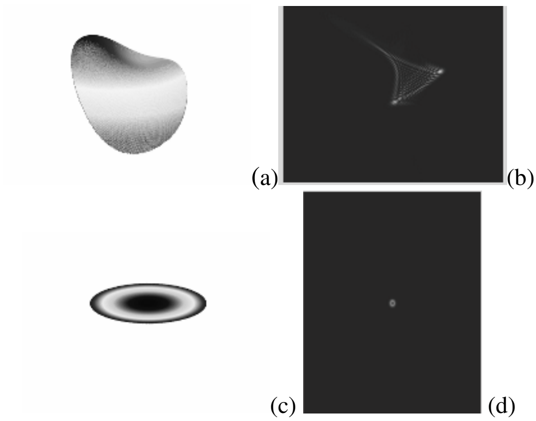


Fig. 2. Wavefront aberration (a,c) and Point Spread Function (b,d) for aberrated and ideal human eye, respectively

3 Deconvolution Overview

Once the PSF is obtained from the measurement of the wavefront aberration function using equation (4), the next step is to find an H^{-1} such that equation (3) is true.

Classical deconvolution defines H^{-1} in frequency as

$$H^{-1}(fx, fy) = \frac{1}{H(fx, fy)} \cdot \tag{6}$$

It is clear that small values of $H(fx,fy)$ will cause singularities in the inverse. Alonso et. all [3] have proposed a method to account for these singularities. Although effective, the parameter selection process for that method is crucial to producing a usable inverse function H^{-1}. These parameters need to be readjusted repeatedly before an acceptable inverse is met.

Additionally, in the model proposed here, noise is not a factor in the deconvolution process, and, therefore, it does not impact the quality of the inverse. We propose a simpler, yet effective form of generating the inverse function, providing regularization for the singularities present in H.

Given a point spread function, H, the inverse, H^{-1} is determined as follows:

$$H^{-1}(fx, fy) = \begin{cases} \dfrac{1}{H(fx, fy)}, & abs(H(fx, fy)) \geq th \\ H(fx, fy), & abs(H(fx, fy)) < th \end{cases} \tag{7}$$

where $0<th<0.1$ specifies a hard threshold to determine the cutoff point for singularity detection.

This threshold yields similar results as in [3], with the added benefit of reducing the sensitivity of the resulting inverse function to the parameter, as well as reducing the problem of parameter choice from three parameters to one.

Unfortunately, the practical implementation of this form of deconvolution necessarily reduces the contrast of *EO* because graphical display devices can only display grayscale values in the range of (0,255). Scaling and shifting of the *EO* is necessary for presentation on modern graphical displays [3], at the cost of the associated loss of contrast.

Although this method of deconvolution yields an adequate inverse, the Fourier domain does not provide a compact form of representing the PSF. Even for the common PSF corresponding to defocus, the associated singularities are spread throughout the frequency domain causing distortion in the form of ringing in the enhanced object, *EO*. This ringing is not necessary to provide adequate enhancement to the user and causes a distortion that is further exacerbated if any attempt at improving the contrast of *EO* is made [1].

The HOWARD algorithm uses wavelet denoising to reduce the presence of the ringing in the enhanced object, *EO*. This has been shown to improve the signal-to-noise ratio (SNR) and mean-square-error (MSE) of signals corrupted by noise [9], [11]. Wavelet image denoising first decomposes an image into l=4, 7, 10…L levels, with wavelet approximation coefficients for each level. In the *Lth* level, the number of wavelet coefficients with a significant amount of energy is typically a subset of all the coefficients calculated, and thus the signal can be reconstructed with an abbreviated number of coefficients [11]. The idea behind denoising is to select the appropriate coefficients to accurately represent the image, while suppressing the noise. In wavelet denoising, one form of selecting the coefficients is by soft thresholding [11], given by

$$w_{soft}(u,v) = \begin{cases} sign(w(u,v))(|w(u,v)| - \tau), & |w(u,v)| > \tau \\ 0, & |w(u,v)| \le \tau \end{cases} \tag{8}$$

where $w_{soft}(u,v)$ are the thresholded wavelet coefficients, $w(u,v)$ are the wavelet coefficients obtained through the decomposition of the image, and τ is the threshold level.

The noise power is assumed to be less than the signal power, thus the noise can be removed from the image. For the enhanced object produced using equation (7), the ringing is assumed to be noise. Thus, for a specified threshold level, the ringing introduced by the regularization is ideally removed from the enhanced object, leaving only the necessary information to satisfy equation (3).

The final step in the algorithm is to apply contrast enhancement to the denoised version of *EO*. HOWARD employs the contrast enhancement method proposed by Alonso et. all [4]. We refer the reader to [4] for a detailed description of the contrast enhancement algorithm. Fig. 3 summarizes the HOWARD algorithm.

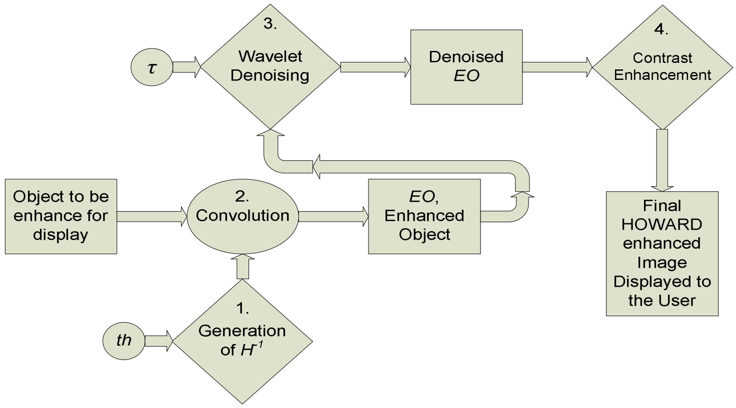


Fig. 3. Summary of the HOWARD algorithm. 1. H^{-1} is generated using the desired threshold, th, and the user's unique wavefront aberration function given in the form of Zernike polynomials. 2. The Object to be enhanced for the user is convolved with H^{-1}. 3. Wavelet denoising is applied to the *EO,* image to produce the denoised version of *EO.* 4. Then contrast enhancement is applied to yield the final, HOWARD enhanced image to be displayed to the user.

4 Considerations for the Verification of the Method

The HOWARD algorithm derives its ability to enhance objects on the graphic display from knowledge of the wavefront aberration function of each user. In order to test the validity of the compensation, an assessment of the image quality perceived by the user is needed to verify the enhancement. Evidently, we do not ordinarily have access to the image that forms on the retina of the user. Thus, in order to provide verification of HOWARD, we expanded upon previous work [2] on developing an artificial eye that can: 1) be measured on a wavefront aberrometer, and 2) provide image capture to verify the processing. The scope of the remaining portion of the paper is dedicated only to this form of validation of HOWARD with example images demonstrating results from the HOWARD algorithm.

A high resolution PixeLINK A782, 6.6 mega-pixel camera was used to capture experimental images through a compound lens composed of a 72mm focal length plano-convex lens, a 10mm adjustable iris, and a 90mm achromat lens, mounted in a standard c-mount lens mount (Edmund Optics). High-Order aberrations were achieved by layering a UV cured Optical Adhesive with a non-uniform surface approximately 0.2mm thick. This created high order aberrations approximately commensurate with those found in some aberrated human eyes. This setup allows images to be captured at 6.6 mega-pixels, while simultaneously providing a measurable, high-order wavefront aberration function.

Lastly, because wavefront aberrometers measure the wavefront aberration function using approximately parallel light rays for viewing conditions [12],[7], the PSF they measure simulates how an object at infinity would appear to the user. Most graphic displays are intended for much shorter distances (~ 50 cm). Thus, an adjustment of the measured Zernike polynomials is necessary to accurately simulate near vision. This is achieved by re-referencing the wavefront from a plane wavefront, to a wavefront that originates at a finite distance from the user.

5 Results

Examples using the digital image of a letter from a standard Early Treatment Diabetic Retinopathy Study (ETDRS) eye test chart are shown in fig. 4. Two display images are used, one un-enhanced, and the other enhanced by HOWARD. Each image is then captured through the high-resolution artificial eye with the high-order wavefront aberration function and corresponding PSF, as shown in fig. 4.

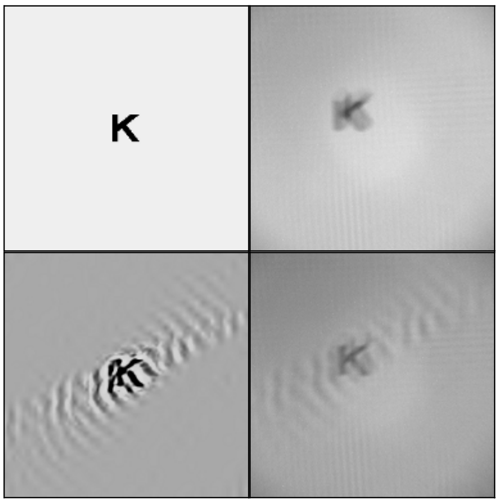


Fig. 4. Top Left- Original Display Object. Bottom Left- HOWARD Enhanced Object. Top Right- Original Display Object captured through the artificial eye. Bottom Right – HOWARD Enhanced Object captured through the artificial eye.

6 Conclusion

In this paper, we propose an efficient method, the High-Order Wavefront Aberration Regularized Deconvolution (HOWARD) algorithm that creates custom enhanced objects for computer users with high-order aberrations in their visual system. The motivation for this stems from the fact that high-order aberrations are not correctable using current ophthalmic methods, making the use of PCs through graphical displays difficult for users with these types of refractive errors. Additionally, existing methods to provide enhancement of vision for high-order wavefront aberrations are costly and inaccessible for the average user.

Upon further development, software-based custom pre-compensation approaches, such as HOWARD, may provide users with otherwise uncorrectable high-order refractory aberrations a way to use GUIs more efficiently in their interaction with computers.

References

1. Alonso Jr., M., Barreto, A., Cremades, J.G. and Jacko, J.A. Image pre-compensation to facilitate computer access for users with refractive errors. Behaviour and Information Technology Journal, 24 (3). 161-173
2. Alonso Jr., M., Barreto, A., Jacko, J.A. and Adjouadi, M., Verification of Computer Display Pre-Compensation for Visual Aberrations in an Artificial Eye. in The Seventh International ACM SIGACCESS Conference on Computers and Accessibility, (Baltimore, MD, 2005), 210-211
3. Alonso Jr., M., Barreto, A.B., Choudhury, M. and Jacko, J.A. Software-based Compensation of Visual Refractive Errors of Computer Users. Biomedical Sciences Instrumentation, 41. 229-234
4. Alonso Jr., M., Barreto, A.B., Jacko, J.A., Choudhury, M. and Adjouadi, M., Improving Computer Interaction for Users with Visual Acuity Deficiencies through Inverse Point Spread Function Processing. in The 2005 IEEE SoutheastCon Conference, (Fort Lauderdale, FL, 2005), 421-427
5. Chou, B. Riding the Wave(front) to Super Vision Review of Optometry Online, 2004
6. Goodman, J.W. Introduction to Fourier Optics. McGraw-Hill Book Company, New York, 1968
7. Liang, J., Grimm, B., Goelz, S. and Bille, J.F. Objective Measurement of wave aberrations of the human eye with the use of a Hartmann-Shack wave-front sensor. Journal of the Optical Society of America, 11. 1949-1957
8. Liang, J., Williams, D.R. and Miller, D.T. Supernormal vision and high-resolution retinal imaging through adaptive optics. Journal of the Optical Society of America, 14. 2884-2892
9. Neelamani, R., Choi, H. and Baraniuk, R. ForWaRD: Fourier-Wavelet Regularized Deconvolution for Ill-Conditioned Systems. IEEE Transactions on Signal Processing, 52 (2). 418-433
10. Parker, J.N. and Parker, P.M. The official patient's sourcebook on Keratoconus: A revised and updated directory for the internet age. Icon Health Publications, San Diego, 2002
11. Strang, G. and Nguyen, T. Wavelets and Filter Banks. Wellesley-Cambridge Press, Wellesley, 1996
12. Thibos, L.N. Formation and Sampling of the Retinal Image. in Valois, K.K.D. ed. Seeing: Handbook of Perception and Cognition, Academic Press, San Diego, 2000, 1-54
13. Wilson, R.G. Fourier Series and Optical Transform Techniques in Contemporary Optics: An Introduction. John Wiley and Sons, Inc., New York, 1995
14. Young, M. AAO preview: A close look at wavefront. EyeWorld, September. 30

“BlindMath” a New Scientific Editor for Blind Students

A. Pepino, C. Freda, F. Ferraro, S. Pagliara, and F. Zanfardino

University of Naples “Federico II” Support Service for Disabled Students (NA) Italy
Via Claudio 21 80125 Naples (IT)
Tel/Fax: +39 081 7683773
pepino@unina.it

Abstract. Today studying mathematics and science for blind students is still an open problem both in Secondary school and at the University in so far as it concerns the effectiveness of teaching approaches, methods and devices (1) Many national and international projects have given a contribution to bring different technical solutions. But, at least in Italy, these solutions are not effective yet in the schools, probably because human, social and organizational aspects have not really been taken into adequate consideration. Here a new method will be proposed, a method which will put forward new technical solutions for editing scientific books and documents but essentially will give adequate answers to operational and practical problems.

1 Introduction: State of the Art in R&D and Application

The national research and outcomes in this field have produced technical solutions deeply correlated with language. We consider, now, the two most important Italian experiences: Braillemat, Lambda.

“Braillemat” (2) is a specialist software designed by a research group belonging to the Computer Centre of the Institute for blind people “Francesco Gavazza”.

This software meets the needs of the primary and secondary school students; it is a reliable and simple system to use but it is based only on the Braille language, which has to be, therefore, currently used both by students and teachers; and this could represent a difficulty for those who do not know it well. One more difficulty concerns some operational problems in exploring the formulas which limit the functionality of the system.

LAMBDA (Linear Access to Mathematics for Braille Device and Audio-Synthesis) (3) is the result of a triennial European Project supported by the European Commission. As the project results are transnational, the system has been based on an absolutely new mathematical code (Lambda Code).

The LAMBDA Code has many close connections with MathML (an XML version for mathematics).

The system interface is based on audio synthesis and an 8 dots Braille code (which is not standard yet). To let the sighted teacher to interact with the class the graphic representation of the formula has been based on a new symbolic notation but this could create some problems to the teacher.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 1171 – 1174, 2006.

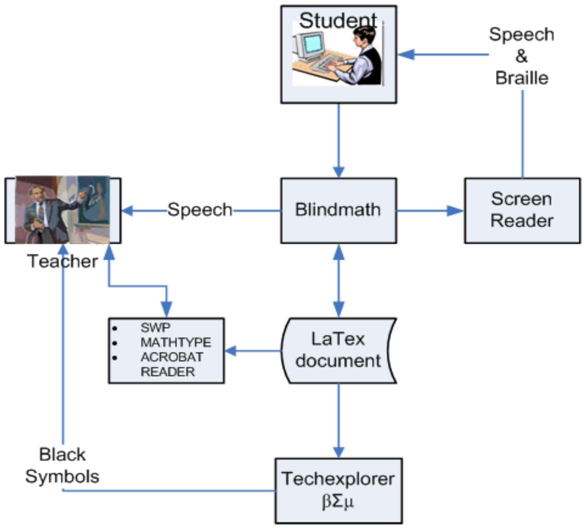


Fig. 1. Blindmath provides speech and Braille to student and teacher, it generates a LaTex document which is reported in symbolic format through techexplore and can be exported in SWP, MathType or pdf

These two technical solutions, which are effective as a matter of principle, are difficult to be used in the real daily school environment.

2 Research and Methodological Approach

2.1 The Technical Approach

The solution here reported starts from the available technical outcomes - voice synthesizer, screen reader technologies, mathematical editors - in order to develop a new "system" able to be more accepted and manageable in the real context. In other words "Blindmath" represents an "integration" of commercial products rather than a new "stand alone" software product. It could be seen as an "open" customizable solution, able to take into account not only the student functional gap but also the teacher operational needs. "Blindmath" system is based on the screen reader software "Jaws" (3) (there are also other software that could be customized for the same goal) commonly used by blind students and on a freeware Mathematical Editor (4) "techexplorer". It includes a code in Delphi environment and a number of jaws customizations (available as shareware) in order to integrate the above mentioned basic tools. This system is based on a LaTex document which is the basic format to translate a symbolic and two-dimensional structure into a linear text row with the necessary structural rules which maintain the univocality of the two-dimension representation.

This system is presented to the student as a specialized scientific digital editor and can be used in any school or at the university even by students who do not know Braille.

Table 1. Activity program

Action	Description
Diffusion	A number of meetings organized by regional district of U.I.C. (Italian Union of Blind) where school managers, teachers, relatives and university researcher discuss about this matter, in order to illustrate how blindmath represents an opportunity for blind students
Preliminary analysis	On the basis of the various requests coming by teachers or_relatives, an accurate analysis of different points of view is made (technical, pedagogical, psychological) in order to verify the real accomplishment of the individual project. An accurate individual project is defined jointly among teachers, school managers, university researchers and relatives.
Agreement	On the basis of the project a technical agreement is undersigned by all the actors involved who take their own responsibility about the project.
Technical solution	All the necessary equipments (PC, Software, Braille display) are prepared and supported by the University thanks to on site and on line technical aid. This guarantees continuity and reliability.
Pre-assessment	Mathematical and general skills are tested through a questionnaire.
Training	Students, Teachers and relatives are trained in order to use blindmath and to provide help and support to the students.
Monitoring	All the activities are regularly monitored
Post-Assessment	Every 3 months mathematical and general skills are tested through a questionnaire

2.2 The Experimental Activity

This system has been deeply tested by a number of blind students of the University of Naples during the different steps of its development. The positive outcome is not altogether satisfactory as it does not guarantee its total adaptability in a real operational context. An useful experimentation should start earlier in the secondary school, in fact, in order both to make blind students able to develop mathematical skills and competencies necessary for scientific faculties, and to test the system in a real environment for at least one year.

This experimentation will be realized in the future, when a number of sub-activities, specifically designed, will be carried out as outlined in the protocol reported in table 1.

3 Preliminary Results

The experimental activity, based on a formal agreement duly signed by the school managers, started in November 2005. It has involved 4 schools; 4 female students from 10 to 15 years have been trained together with their support teachers, 64 training hours in 26 days have been provided up to now. Two students have accomplished the programme and are working with on-demand support only.

All teachers agreed the program. They evaluated the system effective and easy to use. But we are confident to achieve even more satisfactory results in the next months.

References

1. Karshmer, A. I., Bledsoe, C.: Access To Mathematics by blind student http://www.snv.jussieu.fr/inova/villette2002/act5b.htm
2. Braillemat http://www.cisad.it/matematica/braillemat/index.php
3. Lambda Project http://www.lambdaproject.org/ASP/index.aspx?IDMenuAPP=0
4. Jaws software description http://www.freedomscientific.com/fs_products/software_jaws.asp
5. Integre Techexplorer Hypermedia Browser http://www.integretechpub.com/techexplorer/

Translating MathML into Nemeth Braille Code

Paul B. Stanley[1] and Arthur I. Karshmer[2]

[1] MathGenie Project, University of South Florida, USA
pstanley@lakeland.usf.edu
[2] Department of Information Technology, University of South Florida, USA
arthur@lakeland.usf.edu

Abstract. An assistive software application has been created that translates math statements encoded as MathML into Nemeth Braille Code (NBC). This translation is conducted in two phases, the translation of the MathML elements into NBC, then the implementation of rules specific to Nemeth Braille that are irrelevant to MathML. All MathML elements holding semantically relevant information are translated by this program, including the nesting of elements to any level. Some of the syntactical rules inherent to NBC such as the use of the numeric indicator; additional space characters; and some contractions have also been implemented; other rules remain to be incorporated. The NBC can be exported in three ways (1) directly to a refreshable Braille device via a serial connection in real time; (2) saved as a text file then downloaded into a Braille device; and (3) save as a text file then embossed by a third party application. This application allows a person with no Braille experience to enter a math equation into any equation editor that can save that statement in the MathML format, and then convert that statement into Nemeth Braille Code for perusal by the visually impaired.

1 Introduction

The instruction of mathematics is usually conducted with a spoken explanation accompanied by a visual image of the math statement. The spoken version, hereafter referred to simply as the verbal rendering, when accompanied by a visual image tends toward ambiguity. In a classroom structured toward the vision impaired the visual image can be replaced with a tactile representation i.e. an embossed version. A blind student attending a class oriented to sighted students faces the daunting challenge of discerning math statements from their verbal renderings. A notetaker, i.e. a sighted person that takes notes for the blind student, may mitigate this challenge if the notetaker is aware of the vagaries of verbally rendered math otherwise the breakdown in communication merely moves from the teacher to student to notetaker to student.

As an example of the ambiguity intrinsic in the usual verbal rendering of math let's look at this simple math statement:

The square root of x plus 1 over y (1)

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 1175 – 1182, 2006.

This verbal rendering could be perceived in any one of the following ways:

$$\sqrt{\frac{x+1}{y}} \qquad \frac{\sqrt{x}+1}{y} \qquad \sqrt{x}+\frac{1}{y} \qquad \frac{\sqrt{x+1}}{y} \tag{2}$$

Which rendition is correct? It is impossible to eliminate any of the above four formulas based solely on the verbal statement.

Clearly a line of communication is needed between math instructors lacking experience with blind students and their visually impaired charges. This work provides an assistive software tool, usable by anyone who can enter an equation into a math editor, which converts math statements in the MathML format into an unambiguous Braille representation. Input currently requires the use of a third party math editor but requires no experience with Braille.

The scenario of the blind student in the sight oriented classroom was targeted because of its inherent difficulty for the student. If an assistive tool is created for this demanding environment then that tool could likely be applied to less demanding situations.

This work continues the efforts of the MathGenie Project [1,2,3] which has succeeded thus far in converting math in the MathML format to: (1) to a verbal rendering; (2) to an enlargeable visual image for persons with low vision [4]; and (3) a visual image combined with highlighted, spoken text for potential use with the learning disabled [4].

2 Math and Braille Standards

2.1 Braille Code for Representing Math

In 1972 Dr. Abraham Nemeth published the revised version of his text [5] that created a Braille code for math which subsequently became an adopted standard. Nemeth Braille Code is succinct, unambiguous and widely accepted.

2.2 Data-Centricity and MathML

There's a trend in information technology toward separating data from its presentation. The idea is to make the data available but leave the details of its presentation to the application or platform that accesses the data. The Extensible Markup Language (XML) was developed for that purpose for the World Wide Web. XML is a metalanguage used as a template to create other languages. MathML is one of those language spawned from XML. Originally developed to for the inclusion of math in Web pages, MathML has become a standard for representing math statements beyond the Web.

2.3 MathML Language

"MathML is designed to provide the encoding of mathematical information for the bottom, more general layer in two-layer architecture. It is intended to encode complex notational and semantic structure in an explicit, regular, and easy-to-process way for renderers, searching and indexing software, and other mathematical applications." [6]

"It is an XML based language comprised of elements and their attributes. All MathML elements fall into one of three categories: presentation elements, content elements and interface elements" [7].

"This [MathML] project is concerned with only the presentation elements and within that set of elements only those elements encoding the characters and their syntactic and semantic relationships, respectively, the token and constructor elements. Token elements hold the character data, the variables, numbers and operators. The constructors encode the layout schemata, which contain other presentation elements. Each layout schema corresponds to a two-dimensional notational device, such as a superscript or subscript, fraction or table" [8].

2.4 Current Math to Braille Efforts

Tex [9] and Latex [10] are mature protocols utilized for printing math statements. As such, there are various products that translate math statements into Latex [11,12]. At least one commercial product translates text and math into Braille for embossing [13].

This work extends that capacity by converting math in the MathML format directly into Nemeth Braille Code. The user is provided the option of porting the Nemeth Code directly to a refreshable Braille device offering a more dynamic interaction with math, bypassing the embossing process. A new math statement can be made available in the time it takes to create the new statement in a math editor. Additionally, this work has demonstrated the Nemeth Code may be copied to an ASCII text file. This allows downloading Nemeth Code into those Braille devices capable of receiving text files and for exporting to programs that can convert text to Braille for embossing.

A third party math editor [14,15] with the capability to save math statements in the MathML format is still required for the initial input of math to MathML.

3 Research and Methodological Approach

3.1 Overview

This work adheres to the data-centricity concept by storing math in a generic form, then retrieving that math and presenting it in a sensory modality required by the user. The steps for creating the Braille version of a math statement are as follows:

- The math statement is entered into a third party math editor and saved in the MathML format
- The MathML document is opened with the Nemeth translator and validated as well-formed and valid MathML
- The elements of the MathML are translated into ASCII Braille
- Nemeth Rules are applied to create the Nemeth Braille
- The Nemeth Braille is exported
  - To a Braille device in real time
  - To an ASCII text file
    - For downloading into a Braille device
    - For exporting to embossing software tool

3.2 Detailed Discussion

Why begin with MathML format? MathML is the lingua franca of math. It is:

- Standardized
- Nonproprietary and Open source
- There are math editors readily available that have the capacity to save statements in the MathML format [12][13][15]
- Other research efforts in math accessibility utilize MathML; as an interchange format for the translation of math encoded in Latex and various Braille languages [16]; and as a destination format for an optical character reading application that translates scanned text with embedded math formulae into xhtml and MathML [17].

OpenMath is another standardized, XML based protocol for encoding math. Pontelli and Palmer [18] discuss the potential of OpenMath as an interchange format for translating math; Equations encoded in OpenMath can be converted into NBC, however the principle focus of OpenMath is on the content and semantics of math and not its presentation. So for this work, i.e. translating math into the presentation format of Nemeth Braille, OpenMath offers no advantage over MathML-Presentation.

3.2.1 Translating MathML into ASCII Braille

The MathML version of this math statement,

$$\sqrt{\frac{x+1}{y}} \tag{3}$$

is represented in a MathML fragment as:

```
<math>
    <msqrt>
          <mfrac>
                <mrow>
                      <mi>x</mi>
                      <mo>+</mo>
                      <mn>1</mn>
                </mrow>
                <mi>y</mi>
          </mfrac>
    </msqrt>
</math>
```

The strings enclosed in the angled brackets are element tags. The <msqrt></msqrt> tags encode the square root radical. The <mfrac> element encodes the fraction. <mi>, <mo> and <mn> are the token elements holding the character data.

The project application, hereafter referred to simply as the Nemeth application, converts a MathML document as represented by the above diagram into Nemeth Braille Code.

3.2.2 High Level Program Design

Following the Object-Oriented paradigm, classes were created to parse each of the MathML elements, one class for each MathML element. For example, a class called MsqrtParser was designed to parse the msqrt element; the MfracParser was given responsibility for the mfrac element. Each parser object is responsible for collecting the ASCII Braille characters extracted from that construct and its child constructs; instantiating the next parser if required; storing the ASCII Braille string returned to it by the object it spawned; and it in turn return its string to its parent object.

A TextReader object reads and returns the elements and their associated attributes. The translation proceeds like this: The TextReader fetches the next element in the fragment. Referring to the sample MathML above, the math element is fetched first. The MathParser object is instantiated. The MathParser adds no ASCII characters. It does however hold a string data member to store the text returned by the MsqrtParser which it instantiates. The MathParser calls the MsqrtParser into being which in turn instantiates an object of the MfracParser which in turn instantiates an MrowParser. The mrow element encompasses the numerator part of the fraction. MrowParser retrieves the characters x, +, 1 from the token elements mi, mo and mn, returning them to MfracParser. MfracParser retrieves the denominator from the next token element, in this case the character *y*. MfracParser then appends the Braille codes for begin-fraction and end-fraction. It completes its task by returning the string representing the fraction to its parent, the MsqrtParser. MsqrtParser appends the begin-radical and end-radical Braille codes and returns its string to its parent and root element MathParser which in turn passes the completed ASCII Braille string to the GUI for output to the user.

The term ASCII Braille mentioned above requires clarification. Nemeth Braille requires additional characters beyond those gleaned from the MathML tokens. E.g. the question mark signifies the beginning of a fraction. The number sign (#) marks the end of a fraction. The ASCII representation of the Nemeth Code for the fraction

x divided by y becomes: *?x/y#* . (4),(5)

The combination of the string of characters extracted from the math statement and the characters clarifying the construct, the characters indicating the start and finish of the fraction in the above example, is defined here as an ASCII Braille string.

3.2.3 Additional Application of Nemeth Braille Rules

Dr Nemeth fashioned twenty five rules directing the conversion of mathematics into his Braille code [6]. Many of these rules are implemented in the translation of the MathML elements into ASCII Braille. However there are other rules beyond the original translation requiring implementation. The *?* and # indicating the beginning and end of a fraction discussed above is just one example. A few more of the commonly used rules along with brief descriptions are listed here for illustration.

- Numeric Indicator - Indicates transition to or re-affirms continuation of Nemeth Code vis-a-vis literary Braille
- Spacing - Nemeth Code calls for a space character for particular situations e.g. between trigonometric functions and its associated angle; and between a logarithm and its antilogarithm
- Alignment - Embossed Nemeth Code should maintain the same spatial arrangement as it would if presented on a printed page. E.g. the beginning and end characters of the rows in a matrix should vertically align.
- Contractions - Simplified forms of more complex groups of characters. These are not optional in NBC

3.2.4 Outputting to a Braille Device

The final step is passing the Nemeth Braille Code to the user. The development platform used in this project consisted of a laptop computer running Window 2000 and version 7.0 of the JAWS screen reader connected to a BrailleLite M20 via a serial port. The program successfully exported the Nemeth Code via the following methods:

1. The BrailleLite was registered with the JAWS program. Upon sending the Braille string to a text box in the program GUI, the screen reader relayed the Braille string to the BrailleLite.
2. A text file was created from the Nemeth Code, and that file was copied to the BrailleLite.
3. An embossed version was printed by a third party application from the aforementioned text file.

Another plausible scenario not yet implemented is to send the Braille string directly to the Braille device via the BrailleLite's device driver.

4 Results

The MathMLtoNemeth application successfully translates all relevant MathML elements into Nemeth Braille Code. However, as of this writing, not all of the additional NBC rules have been implemented so the output is not yet fully compliant with Nemeth Braille.

Two methods were used to transfer the Nemeth Code to the Braille device, 1) synchronously via a serial port connection and 2) exported to the Braille device as a text file.

This application allows a person with no Braille experience to enter a math statement into any equation editor that can save that statement in the MathML format, and then convert that statement into Nemeth Braille Code for perusal by the visually impaired.

5 Future Work

This application is one node in a web of communication not yet fully formed. The following areas await future implementation.

- Convert print materials to MathML. While it's possible for a person with no Braille experience to create Braille representations of math statements, it's still burdensome for an instructor to duplicate an entire course of lecture material and exams one formula at a time.
- Create a math editor for the visually impaired. This editor will open a line of communication from the blind user to anyone with an application that can read a MathML document [19], i.e. the sighted recipient. The editor should allow users to
    - Open other peoples equations encoded as MathML fragments
    - Create math statements and save them in the MathML format
    - Allow the user to manipulate equations in real time
- Eliminate the screenreader intermediary and connect directly with the Braille devices through device drivers

Acknowledgements

The work reported in this paper was funded through a grant from the U.S. Department of Education, National Institute fro Disability and Rehabilitation Research (NIDRR) under Grant Number H133G010046.

References

1. Karshmer, A.I., Bledsoe, C. and Stanley, P.B.: "The Architecture of a Comprehensive Equation Browser for the Print Impaired," 2005 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC'05), pp. 319-320
2. Karshmer, A.I: "Mathematics for Blind People," Introduction to the Special Thematic Session, Computers Helping People with Special Needs, 9th International Conference, ICCHP 2004, Paris, France, July 7-9, 2004, Proceedings. Lecture Notes in Computer Science 3118 Springer 2004, ISBN 3-540-22334-7, pp 614-619
3. Karshmer, A.I. and Bledsoe, C.: "Access to Mathematics by Blind Students," Introduction to the Special Thematic Session, Computers Helping People with Special Needs, 8th International Conference, ICCHP 2002, Linz, Austria, July 15-20, Proceedings. Lecture Notes in Computer Science 2398 Springer 2002, ISBN 3-540-43904-8, p 471
4. Stanley, P.B., Bledsoe, C. and Karshmer, A.I.: "Utilizing Scalable Vector Graphics in the Instruction of Mathematics to the Print Impaired Student," Computers Helping People with Special Needs, 9th International Conference, ICCHP 2004, Paris, France, July 7-9, 2004, Proceedings. Lecture Notes in Computer Science 3118 Springer 2004, ISBN 3-540-22334-7, pp 626-629
5. Nemeth, A.: "The Nemeth Braille Code for Mathematics and Science Notation 1972 Revision," American Printing House for the Blind, 1972
6. w3c. "MathML 2.0 Implementation and Interoperability Report," `http://www.w3.org/Math/iandi/`, Accessed February 15, 2006
7. w3c. "Introduction to MathML - Overview: Mathematical Markup Language (MathML) Version 2.0 (Second Edition)," Accessed February 15, 2006
8. w3c. "Mathematical Markup Language (MathML) Version 2.0 (Second Edition) 2.1.1 Taxonomy of MathML Elements," `http://www.w3.org/TR/2003/REC-MathML2-20031021/chapter2.html#id.2.1.1`, Accessed February 15, 2006

9. Knuth, D.: “TeXbook,” Addison-Wesley Professional, 1986
10. Lamport, L.: “LaTex: A Document Preparation System, 2nd Edition”, Addison-Wesley Professional, 1984
11. Scientific Workplace, Scientific Word, MacKichen, `http://www.mackichan.com/`
12. Mathematica, Wolfram Research, `http://search.wolfram.com/`
13. Duxbury Braille Translator, Duxbury Systems, Inc.,
14. `http://www.duxburysystems.com/`
15. Amaya XML Editor, W3C, `http://www.w3.org/Amaya/`
16. MathType Equation Editor, Design Science, Inc., `http://www.dessci.com/en/products/mathtype/`
17. Archambault, D., Fitzpatrick, D., Gupta, G., Karshmer, A., Miesenberger, K. and Pontelli, E.: “Towards a Universal Maths Conversion Library”, Computers Helping People with Special Needs, 9th International Conference, ICCHP 2004, Paris, France, July 7-9, 2004, Proceedings. Lecture Notes in Computer Science 3118 Springer 2004, ISBN 3-540-22334-7, pp 664-669
18. Suzuki, M., Kanahori, T., Ohtake, N. and Yamaguchi, K.: “An Integrated Software for Mathematical Documents and Its Output with Accessibility”, Computers Helping People with Special Needs, 9th International Conference, ICCHP 2004, Paris, France, July 7-9, 2004, Proceedings. Lecture Notes in Computer Science 3118 Springer 2004, ISBN 3-540-22334-7, pp 648-655
19. Pontelli, E. and Palmer, B., Translating Between Formats for Mathematics: Current Approach and an Agenda for Future Developments, Computers Helping People with Special Needs, 9th International Conference, ICCHP 2004, Paris, France, July 7-9, 2004, Proceedings. Lecture Notes in Computer Science 3118 Springer 2004, ISBN 3-540-22334-7, pp 620--625
20. The Firefox Web Browser displays MathML documents innately. MathPlayer by Design Science is a free plug-in for Internet Explorer. These are just two examples.

New Environment for Visually Disabled Students to Access Scientific Information by Combining Speech Interface and Tactile Graphics

Toshihiko Komada[1], Katsuhito Yamaguchi[1], Fukashi Kawane[1], and Masakazu Suzuki[2]

[1] Junior College Funabashi Campus, Nihon University, 7-24-1 Narashinodai, Funabashi, Chiba 274-8501, Japan
komada@maphy.jcn.nihon-u.ac.jp, eugene@gaea.jcn.nihon-u.ac.jp, kawane@gaea.jcn.nihon-u.ac.jp
[2] Faculty of Mathematics, Kyushu University, 6-10-1 Hakozaki, Higashi-Ku, Fukuoka 812-8521, Japan
suzuki@math.kyushu-u.ac.jp

Abstract. We recently developed a math document editor with a function of speech output. Using this software named "ChattyInfty," visually disabled students not only can read scientific documents including math expressions with speech output but also can author or edit them. One of remarkable features of ChattyInfty is its availability of recognition results by our math OCR system. By combining them, we construct a system to enable students with visual disabilities to access scientific documents in print for themselves. Furthermore, using tactile graphic tools, we realize an environment in which they can grasp and correct recognition errors without any aid of sighted people.

1 Introduction

The development of information technology remarkably improves an environment for visually disabled people to access printed materials. However, as concerns scientific information, it is still very hard for them to read or to edit it for themselves. Ordinary OCR (optical character recognition) systems and screen readers cannot treat such information since it includes many special symbols and notations such as complicated equations.

For this decade, various approaches have been tried to improve math accessibility. In 1994, T. V. Raman reported on his "Audio System for Technical Readings" that could read out math documents in LaTeX format [1]. "EMACSpeak project" currently succeeds to his work [2]. Recently, many research groups are working on new approaches based on MathML. For instance, "Mathplayer" developed by Design Science is a plug-in for the Internet Explorer to read out embedded math expressions in MathML [3]. The Lambda Project also studies Braille and speech outputs of math documents in various European languages;

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 1183–1190, 2006.

a prototype version of their linear math editor, “the Lambda editor” is released [4]. “iGroup UMA (Universal Math Access)” is developing automatic translators between various Braille-based math notations [5]. Some commercial Braille translators can produce math documents in various Braille notations [6,7].

These works make great contributions to producing accessible scientific materials for visually disabled students; however, there still remain unsolved problems that we have to work on. At first, although students with print disabilities become able to read given (accessible) information, it is still quite hard for them to author or to manipulate it. Furthermore, they have no way to make math documents in print accessible for themselves. They cannot easily share same scientific materials with their sighted classmates or instructors.

Our research project named “InftyProject” [8] has been doing research on digitization and computerized processing of scientific information and its accessibility. As were reported in the ICCHP2000 [9] and the ICCHP2004 [10], for instance, we developed the OCR system for mathematical documents and prototype Braille translator and a speech-output system for recognition results. It enabled visually disabled students to access mathematical information for themselves up to a certain level. Improving the speech interface, we lately developed an accessible math-document editor with a function of speech output named “ChattyInfty.”

As was mentioned above, recent progress of the OCR technology remarkably improves information environment for persons with print disabilities; however, most of OCR software cannot treat printed scientific documents. On the other hand, our math OCR system named “InftyReader” [8,9,10,11] can recognize them correctly and convert the recognized results to various accessible formats. Its extended version, “InftyReaderPlus” can recognize math documents in PDF (Adobe Portable Document Format) as well.

One of remarkable features of ChattyInfty is its availability for recognition results by InftyReader (or InftyReaderPlus). Actually, combining them, we construct a system to enable visually disabled students to access scientific documents in print for themselves. Furthermore, by making use of tactile graphic tools, we realize a new environment in which they can understand original scanned images as tactile graphics.

Here, at first, the specifications of ChattyInfty are shown in the Section 2. Next, in the Section 3, the new environment that is realized by combining ChattyInfty and InftyReader is discussed. In the Section 4, we evaluate our system for some examples. Finally, the conclusion is discussed in the Section 5.

2 Specifications of ChattyInfty

In InftyProject, we have originally developed a math document editor named “InftyEditor” [8] to manipulate the recognition results by InftyReader. InftyReader initially outputs the results as files in “IML (Infty Markup Language)” format peculiar to Infty software, which is essentially XML. We can convert them to ones in other various formats such as LaTeX, MathML and so on via

InftyEditor. Furthermore, given files in those formats can be also imported directly into InftyEditor. If InftyEditor were accessible, this environment of scientific information processing would be available for persons with print disabilities. In order to achieve that, we worked on incorporating a function of speech output into it.

First of all, based on standard ones in the math education of USA, we defined a manner of aloud reading for each mathematical symbols and special notations such as superscripts, subscripts, fractions, radicals and so on (we call them "the math syntax"). When assigning the manner of aloud reading, we also referred to speech output by Mathplayer [3]. Using Microsoft Speech API, we developed own speech interface for contents displayed on the main window and for the context menu in which math symbols and special notations are listed. After several improvements for about two years, this software is nearly completed and named ChattyInfty recently.

Information on the main window of ChattyInfty is completely equivalent to one of InftyEditor. All of math expressions are displayed in the ordinary print style. ChattyInfty reads aloud not only texts but also those math expressions. Visually disabled students, therefore, can easily share the same materials with their classmates and instructors.

It should be noted, however, that they need to use a screen reader such as JAWS [12] along with ChattyInfty if they hope to operate ChattyInfty with the aid of speech output only. It reads out just main contents and the list of symbols in the context menu with its own speech interface; other parts such as pull-down menus are not read aloud. When ChattyInfty is used along with the screen reader, some conflicts between their speech outputs may take place. Actually, we have already confirmed such a trouble between JAWS and ChattyInfty. In order to avoid it, we prepared JAWS's script files for ChattyInfty. We believe that such treatments are also possible for other popular screen readers.

In a text area, operations for browsing with aloud reading are almost same as ones in popular screen readers. Using arrow keys, we can also browse math expressions; however, in math areas, a behavior of cursor with pressing the arrow keys is a little different from the ordinary. For instance, pressing the up arrow key moves the cursor to the superscript position of each symbol. In order to move it to the previous line, we have to press Ctrl + the up arrow key. Pressing the right arrow key from the top of fraction moves the cursor from the numerator to the denominator. Users are required to understand properly those features.

In ChattyInfty, one of two different types of aloud reading: Plain reading mode and Detailed reading mode can be selected. In the detailed reading mode, attributive information concerning each character such as "capital" is read aloud. Pressing Ctrl + F5 key switches these two modes, alternately. Even in the plain reading mode, one can confirm the detailed reading of each character at the cursor position by pressing Ctrl + Alt + H. Using "Reading Table Editor," users can change a manner of aloud reading for each symbol. The "Default" button cancels all of the manner that the user has defined up to that time.

As concerns authoring, math expressions ranging in level from math textbooks of the secondary school to higher technical papers can easily written with speech guide only. One can input a variety of special symbols and special notations from the Context menu that pops up by pressing the Application key. Since mathematical symbols are classified into several categories such as Greek letters, Relative operators, Functions and so on, we can smoothly find necessary one.

We can customize entry listed in the Context menu. “Math level” in the Setting menu provides the following 4 levels:

Junior high school math, Senior high school math,
University math, Professional math.

Since ChattyInfty prepares enough kinds of symbols for highly technical materials, if the appropriate setting were not chosen, it might be a little hard for us to find a necessary symbol with speech guide only.

For the time being, mathematical information on ChattyInfty cannot be converted directly to Braille. However, since IML files can be changed to files in LaTeX, MathML and so on, one can translate them into Braille by using other Braille translators. Actually, we confirmed that the Duxbury Braille Translator (DBT) [6] can almost properly convert LaTeX files produced by ChattyInfty to Nemeth and UEB (Unified English Braille) code. Translation into British and French Braille math code must be also possible with DBT.

3 New Environment to Access Scientific Information in Print

Our math OCR system, InftyReader can recognize only binary (black and white) page images carefully scanned in either 600 DPI or 400 DPI. The image files have to be prepared in either TIFF, GIF or PNG format. Its extended version, InftyReaderPlus accepts PDF files as well. Since the recognized results can be imported directly into ChattyInfty, one not only can access but also can edit them easily with speech output.

However, InftyReader, itself is not necessarily accessible. We, therefore, developed a new user interface for visually disabled students to operate InftyReader. When ChattyInfty is installed, this interface becomes available. By selecting an image file in “My Computer” window and by pressing the application key, one can perform the command, “Recognize with Infty System.” Then, the OCR runs without any additional operations. After the recognition is completed, ChattyInfty will be opened automatically, and the result will be displayed on its main window. Incidentally, one can also perform OCR from the File menu of ChattyInfty.

Even if the OCR were highly accurate, recognition errors could take place at a certain rate. In the ordinary OCR software, most of those errors can be found and corrected by so-called “a spell checker. ” As concerns math expressions, however, it is quite difficult to grasp them from the context of the recognized result only. InftyReader holds information of 2-dimensional coordinates set on the scanned image even after the recognition was completed. When the recognized result is displayed on the main frame of ChattyInfty, the "Sheet View window" pops up to show the original page image. Since the image that exactly corresponds to an area around the cursor position on the recognized result is displayed, sighted persons can easily check and correct errors of the recognition.

It is remarkable that graphic information on the Sheet View window is also available for visually disabled persons. By pressing F4 key, the focus moves to the Sheet View window. If a certain tactile-graphic tool such as View Plus embossers [7] is available, one can output the image directly from this window as tactile graphics. Therefore, even if users were totally blind, they could grasp and correct recognition errors for themselves with the tactile graphics.

The Sheet View window has its own setting menu. One can reduce or magnify displayed images ranging in size from 10% to 1000%. Since the size of tactile graphics is linked with this setting, one can easily tune up its resolution.

As concerns math education for visually disabled students, they do not necessarily need to have a knowledge on printed math expressions. They, of course, can study mathematics using Braille and aloud reading. Such knowledge, however, sometimes help them with communicating mathematical information to their sighted classmates and instructors. Using InftyReader and ChattyInfty, they can easily understand real forms of math symbols and formulas in print. It may play a certain role in math education.

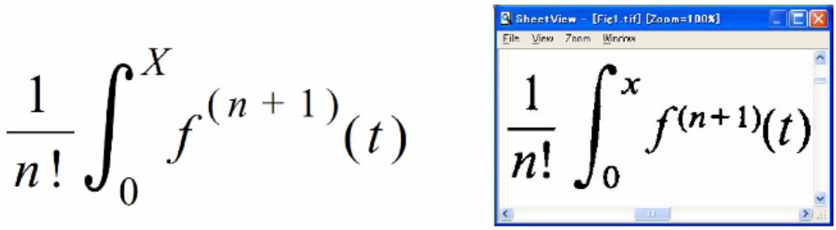
The extended version of InftyReader, InftyReaderPlus can recognize scientific documents in PDF. Many electronic documents such as on-line journals are currently provided as PDF files. Combining InftyReaderPlus and ChattyInfty, students with visual disabilities can also read them without the aid of sighted people.

4 Evaluation of the System

We tested our system to evaluate if visually disabled students can grasp correctly math information produced by it. Scanned images taken from several College textbooks such as Algebra, Calculus, Complex Analysis and so forth were recognized. Using no visual information, we compared and verified their tactile graphics with the recognized results.

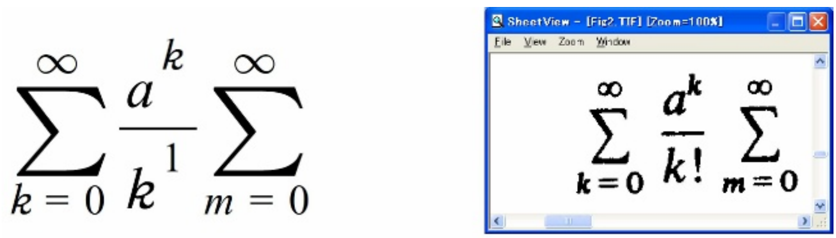
In this test, we used two kinds of tactile graphic tools. One was "Dot View 2" [13] by K.G.S. which has refreshable 48 by 32 dots display. It can display graphical information as shown on the computer screen in real time. When the Sheet View window is focused, the image on it is automatically displayed on the Dot View 2. The other was "MAX Braille embosser" by View Plus [7]. "Tiger Software Suits" bundled it automatically converts colored and shaded images to 20DPI 3-D tactile graphics and outputs it to the embosser in the same manner as ordinary Windows printer drivers. The image on the Sheet View window can be embossed directly from the own Print menu of this special window.

Fig.1a and 1b are a recognized result and its original image taken from a textbook of the calculus, respectively. In this case, since superscripts, subscripts, a fraction, a factorial mark and so on are all properly recognized, students with visual disabilities can understand this math expression with the recognized result only.



a) Recognized result b) Original image

Fig. 1



a) Recognized result b) Original image

Fig. 2

a) Recognized result b) Original image

Fig. 3

Fig.2a and 2b show a recognized result and its original image taken from a textbook of the complex analysis, respectively. In this case, although most of symbols and all special notations are correctly recognized, only the factorial mark is misrecognized as the power 1. Since one can easily grasp this error with both of the Dot View and the MAX embosser, there is no problem for visually disabled students to understand this equation with speech output and tactile graphics.

Fig.3a and 3b are a recognized result and its original image taken from a textbook of the partial differential equation, respectively. In this case, although most of symbols and a special notation (a matrix) are correctly recognized, only

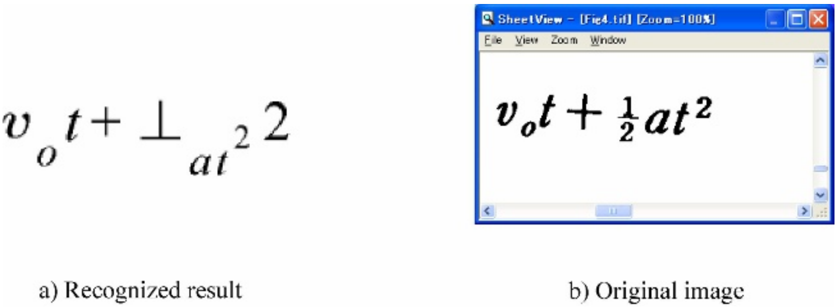


Fig. 4

two of the letter 1 are misrecognized as the letter l. These errors also can be easily realized with both of the Dot View and the MAX embosser.

Fig.4a and 4b are a recognized result and its original image taken from a textbook of the physics, respectively. In this case, since the quality of image is not good, a major misrecognition takes place. Since a fraction bar touches the numerator 1, they are recognized as a single "is perpendicular to" sign. Letters after fraction are interpreted as a subscript to that sign. The denominator 2 are recognized as an independent letter to them. In this case, it is very difficult to understand this expression with the recognized result only; however, one can also grasp these errors with both of the Dot View and the MAX embosser.

Since tactile graphics by View Plus embosser have higher resolution, they seem more favorable to realize original images. When using the Dot View 2 in this test, we needed to zoom in images on the Sheet View window, at least, up to 300 or 400%. The advantage of Dot View 2 is its better efficiency. Printing out by the Braille embosser requires a lot of time and costs.

It becomes clear anyhow that, in most cases, visually disabled students can grasp original scanned images as tactile graphics if necessary. Provided that they understood original forms of printed math expressions properly, they could easily correct the recognition errors. It is shown that they can convert highly technical math textbooks smoothly to accessible ones for themselves using our system.

5 Conclusion

Combining ChattyInfty and InftyReader, we realize the new environment in which students with visual disabilities not only can access but also can author scientific documents including mathematical expressions. One of its remarkable features is that both of original scanned images and their recognition results are available for them. It must play an important role in the math education since our system enables them to share same materials with sighted classmates and instructors. Furthermore, using our system, both of persons with and without visual disabilities, for themselves, can convert scientific documents in print to various accessible formats such as LaTeX, MathML besides IML. This function can be also combined with other many assistive tools to produce accessible materials. Infty software can be downloaded from our web site [8].

We are now developing Braille-output interface for ChattyInfty. We also need to develop an interface to enable students with visual disabilities to access smoothly tables and figures included in math documents. We intend to improve it so that it will become more user-friendly and useful.

References

1. Raman, T. V.: "Audio System for Technical Readings". Dissertation for Cornell University (1994).
2. http://emacspeak.sourceforge.net/
3. http://www.dessci.com/mathplayer/
4. http://www.lambdaproject.org/
5. http://karshmer.lklnd.usf.edu/˜igroupuma/
6. http://www.DuxburySystems.com/
7. http://www.ViewPlus/products/
8. http://www.inftyproject.org/en/
9. Fukuda, R., Ohtake, N., Suzuki, M.: "Optical recognition and Braille Transcription of Mathematical Documents". Proceedings of Computers Helping people with Special Needs, 7th International Conference ICCHP2000, Karlsruhe (2000) 711-718.
10. Suzuki, M., Kanahori, T., Ohtake, N., Yamaguchi, K.: "An Integrated OCR Software for mathematical Documents and Its Output with Accessibility," Computers Helping people with Special Needs, 9th International Conference ICCHP2004, Paris, July 2004, Lecture Notes in Computer Sciences 3119, Springer (2004) 648-655.
11. Suzuki, M., Tamari, F., Fukuda, R., Uchida, S., Kanahori, T.: "INFTY - An integrated OCR system for mathematical documents,". Proceedings of the 2003 ACM Symposium on Document Engineering (DocEng2003), Grenoble, Ed. C.Vanoirbeek, C.Roisin, E. Munson (2003) 95-104.
12. http://www.freedomscientific.com/
13. http://www.kgs-jpn.co.jp/

Canonical MathML to Simplify Conversion of MathML to Braille Mathematical Notations

Dominique Archambault and Victor Moço

Université Pierre et Marie Curie, Inserm U483/INOVA,
9 quai Saint Bernard, 75252 Paris cedex 5, France

Abstract. This paper describes the Canonical MathML, a tentative to unify MathML structures in a deterministic way in order to simplify transcription into Braille. All Mathematical structures that are necessary to perform a correct transcription into Mathematical Braille are recognised and rewritten in a unique way. Additionally Canonical MathML is valid MathML so it can be used with common tools which handle MathML. The Canonical MathML was successfully used to build several transcribers from MathML to Braille national codes.

1 Introduction

Writing and reading Mathematical formulas has always been a difficult task for Blind people. Indeed sighted people use a graphical representation, in two dimensions, in which the layout helps the reader to understand the meaning of a formula. People with visual impairment use linear modalities, like Braille or speech synthesis, and therefore don't have access to this powerful help.

One of the main problems they encounter is the length of formulas, which tends to increase dramatically with the complexity of the formulas represented. In order to reduce this problem, the inventors of specific Braille notations to represent mathematical formulas used various techniques, including contextual grammars [1,2].

These notations are well suited to the use by Braille readers, but their production has always been a problem. The main reason is that there are various such notations, corresponding to various languages, and the various location where they were developed. Then when a document is transcribed and exists in a digital format, it's very difficult to exchange with readers from another language (even though they can read the plain text in the other language itself without difficult). The same difficulty occurs with powerful software tools which have been developed in various laboratories and companies during the last two decades to produce Mathematical Braille (Labradoor (LAtex to BRAille DOOR) [3] produces Marburg Braille, Insight [4] handles Nemeth Braille, Bramanet works with French: http://handy.univ-lyon1.fr/projets/bramanet). Additionally some other very interesting projects handle MathML formulas, like the Math Genie [5] and Infty [6]. These projects would benefit of an easy way to access to a Braille output.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 1191–1198, 2006.

Today more and more digital documents are available via the Web, or in digital libraries exists [7]. MathML, maintained by the W3C Consortium became very quickly the standard for representation of mathematical contents in digital documents. Then providing converters from MathML to Braille notations will make life much easier for blind scientists and science apprentices.

In order to propose to users from different mathematical Braille cultures access to more documents, and access to a great variety of software tools, the International Group for Universal Maths Access started to develop a Universal Maths Conversion library [8]. This library is intended to provide conversion functionalities to and from various Braille notations, as well as standard representations, like MathML. Actually MathML was chosen by the members of the group to be the central representation.

The problem with MathML is that it allows to represent the same mathematical contents in many ways and then the conversion into Braille is not that easy. Furthermore Braille notations, if they are quite different from each other, at a certain level, share also a lot of characteristics. That's why we started to develop a set of tools to unify the way each mathematical structure is to be represented in MathML in order to simplify transcription into Braille. Actually 90% of the transcription work is done generically on the MathML representation, and this work is the same for any Braille code. Then the last part of the work, the transcription into Braille itself, is to implement the specific grammatical rules of the target notation, and to turn every mathematical symbol into its Braille representation, using a dictionary.

The intermediate MathML representation is specified so each Mathematical structure can be represented in a unique way, so the transcription into Braille is deterministic. In other words, all MathML structures that would be represented in the same way in Braille are unified into a unique MathML representation. We call this representation Canonical MathML. One important point about our Canonical MathML is that any Canonical MathML formula must be valid MathML in all cases.

2 Canonical MathML Specification

The next section presents the Canonical MathML structure, in different cases. Description of other structures (like limits) can be found on the umcl-demo website (see conclusion).

2.1 Subscripts and Superscripts

Subscript (respectively superscripts) must be represented using the `msub` (respectively `msup`) element. The element `msub` (resp. `msup`) must have 2 child nodes. None of them can be empty.

Subscripts with superscript must be represented using the element `msubsup`. It must have 3 child nodes. None of them can be empty.

The following table shows the Canonical MathML code to represent structures from these 3 cases:

Subscript	Superscript	Subscript with superscript

<msub>
  child1
  child2
</msub>
	<msup>
	  child1
	  child3
	</msup>
		<msubsup>
		  child1
		  child2
		  child3
		</msubsup>

Where: – `child1` is a term, – `child2` is an indice and – `child3` is an exponent. Any of them can be simple elements (e.g. `<mi>a</mi>`) or complex elements encapsulated in a `<mrow>` element (This applies to all terms called `child`i in the rest of the paper, except when specified).		

Examples:

a_i	a^{i+1}
<msub> <mi>a</mi> <mi>i</mi> </msub>	<msup> <mi>a</mi> <mrow> <mi>i</mi> <mo>+</mo> <mn>1</mn> </mrow> </msup>

2.2 Parenthesis Groups

We call "parenthesis group" a group of elements delimited by an opening tag and closing tag. Then absolute values and intervals are considered as parenthesis groups. In MathML, there are several ways to represent parenthesis groups. The specific tag `<mfence>` allows to specify the opening and closing tags. But this tag is barely used by the tools which produce MathML. Generally, the group delimited by parenthesis are not explicitly defined on the MathML notation.

For instance, the formula $(a + b)$ is often represented in MathML as in the left column of the following table. In the canonical notation, all the group must be included in a `<mrow>` tag, as in the right column. Then the content of the parenthesis group must be included in a `<mrow>` too. This allows to differentiate algebraic parenthesis from intervals (see in the second example the representation of the interval $[0, +\infty[$).

NB: in the next examples, both representations are valid according to the MathML specification. There are a lot of ways to represent them, but only one in Canonical MathML.

$(a+b)$	
Valid in MathML	Canonical MathML
<mo>(</mo> <mi>a</mi> <mo>+</mo> <mi>b</mi> <mo>)</mo>	<mrow> <mo>(</mo> <mrow> <mi>a</mi> <mo>+</mo> <mi>b</mi> </mrow> <mo>)</mo> </mrow>

$[0,+\infty[$	
Valid in MathML	Canonical MathML
<mo>[</mo> <mn>0</mn> <mo>,</mo> <mo>+</mo> <mi>∞</mi> <mo>[</mo>	<mrow> <mo>[</mo> <mn>0</mn> <mo>,</mo> <mrow> <mo>+</mo> <mi>∞</mi> </mrow> <mo>[</mo> </mrow>

This representation makes it easier to navigate in mathematical formulas, as well as to translate formulas into mathematical braille. It is necessary to understand that converting a formula into canonical MathML does not modify the visual representation of formulas.

In the canonical MathML, a parenthesis is always represented as follows :

```
<mrow>
   <mo>opening symbol</mo>
   list of child nodes
   <mo>closing symbol</mo>
</mrow>
```

The list of child nodes can be:

- empty (e.g. $f()$);
- a simple element or a complex element included in a `mrow` in the case of algebraic parenthesis;
- a list of simple or complex elements in the case of intervals, where separators are simple `mo` elements.

2.3 Fractions

Fractions must be represented using the element `mfrac`. It must have 2 child nodes, none of them can be empty. The first child node correspond to the numerator and the second to the denominator.

The following table shows the Canonical MathML code to represent a fraction:

Fraction	Where:

<mfrac>
  child1
  child2
</mfrac>
	– child1 is the numerator,
	– child2 is the denominator.

Examples:

$\frac{1}{x}$	$\frac{1}{x+1}$	$\frac{x-1}{x+1}$
<mfrac> <mn>1</mn> <mi>x</mi> </mfrac>	<mfrac> <mn>1</mn> <mrow> <mn>x</mn> <mo>+</mo> <mi>1</mi> </mrow> </mfrac>	<mfrac> <mrow> <mn>x</mn> <mo>-</mo> <mi>1</mi> </mrow> <mrow> <mn>x</mn> <mo>+</mo> <mi>1</mi> </mrow> </mfrac>

2.4 Square Root, Roots

Square roots must be represented using the element `msqrt`. It must have only 1 child node, which cannot be empty.

Roots must be represented using the element `mroot`. It must have 2 child nodes, which cannot be empty. The first child node is the radicand and the second one the index.

The following table shows the Canonical MathML code to represent a square roots and roots:

Square root	Root

<msqrt>
  child1
</msqrt>
	<mroot>
	  child1
	  child2
	</mroot>

Where: – child1 is the radicant, – child2 is the index (warning: the index is the second one).	

2.5 Summations, Products and Integrals

Summations, Products and Integrals must be represented using a <mrow> with 2 child nodes. The first child node is representing the opening of the structure. It may be a single element or a <mrow> of 2 or 3 child nodes, depending if the initialisation and limit are present. In any case the symbol representing the summation or product symbol must be there (basically Σ, Π or $\int$). The second one is the content of the summation or product. If a factor dx is present it must be included in the second child.

The following table shows the Canonical MathML code to represent summations in 3 cases:

Whole structure	limit is not there	Minimum structure
<mrow> <msubsup> child1 child2 child3 <msubsup> child4 </mrow>	<mrow> <msub> child1 child2 <msub> child4 </mrow>	<mrow> child1 child4 </mrow>
Where: – child1 contains the summation or product symbol (Σ, Π or $\int$...), it **must** be a simple element. – child2 is the bottom element (initialisation), – child3 is the top element (limit), – child4 is the content of the summation or the product.		

Examples:

$\Pi_i a_i$	$\int_{x=0}^{\infty} f(x)$
<mrow> <msub> <mi>Π</mi> <mi>i</mi> </msub> <msub> <mi>a</mi> <mi>i</mi> </msub> </mrow>	<mrow> <msubsup> <mo>∫</mo> <mrow> <mi>x</mi> <mo>=</mo> <mn>0</mn> </mrow> <mn>∞</mn> </msubsup> <mrow> <mi>f</mi> <mo>⁢</mo> <mrow> <mo>(</mo> <mi>x</mi> <mo>)</mo> </mrow> </mrow> </mrow>

3 Production of Canonical MathML

We have developed a set of 8 XSLT style sheets which allow to produce canonical MathML from any MathML source.

- `01_prepare.xsl:` process `<mrow>` tags (remove the ones that are not necessary. Save the ones which delimit complex structures, Sums, etc...
- `02_group1.xsl and 03_group2.xsl:` process opening and closing tags (like in parenthesis groups). NB: Currently only structures with an opening tag **and** a closing tag are supported.
- `04_unary_op.xsl:` search for unary operators and insert them into `<mrow>` when necessary. Add "InvisibleTimes" when necessary.
- `05_structure.xsl:` define XML structure for various specific MathML subtrees (e.g. sums, products, integrals).
- `06_binary_op.xsl:` process binary operators. This style sheet must be called 3 times successively with three different levels of operators.
- `07_canonical_mathml.xsl:` remove some temporary XML tags used by the other stylesheets that are not valid in MathML to obtain valid MathML.
- `08_add_index.xsl:` insert IDs to all MathML tags.

4 Conclusion

To produce Braille from Canonical MathML, we need to develop another XSLT stylesheet to implement the grammatical rules of the target Braille code, and a dictionary of symbols. A set of software scripts were developed to make it easier the development of the dictionary, using OpenOffice.org spreadsheet.

Currently the Canonical MathML and the corresponding XSLT style sheet allow to process any formula in algebra up to the end of secondary school. Implementation of matrices and limits is ongoing but should be available at the camera ready deadline.

Output style sheets and dictionaries were developed for French (2 notations) and Italian Braille codes. Marburg and British are under development in collaboration with universities of Linz and Dublin (DCU).

Additionally a portable C library, UMCL (Universal Math Conversion Library), was developed in order to make it easier the use of the converters. UMCL is an open-source project. A demonstration website gives access to documentation about the Canonical MathML, to an online conversion tool based on UMCL and to the source code:

`http://inova.snv.jussieu.fr/umcl-demo`

References

1. Archambault, D., Batušić, M., Berger, F., Fitzpatrick, D., Miesenberger, K., Moço, V., Stöger, B.: The universal maths conversion library: an attempt to build an open software library to convert mathematical contents in various formats. In Stephanidis, C., ed.: Proceedings of 3nd International Conference on Universal Access in Human-Computer Interaction (joint with HCI International 2005), Las Vegas, Nevada, USA (2005) 9 pages (proceedings on CD-Rom).

2. Stöger, B., Miesenberger, K., Batušić, M.: Multimodal interaction with mathematical formulae: Access, communication and support for blind people in doing mathematics. In Stephanidis, C., ed.: Proceedings of 3nd International Conference on Universal Access in Human-Computer Interaction (joint with HCI International 2005), Las Vegas, Nevada, USA (2005) 8 pages (proceedings on CD-Rom).
3. Batusic, M., Miesenberger, K., Stoeger, B.: Labradoor, a contribution to making mathematics accessible for the blind. In Edwards, A., Arato, A., Zagler, W., eds.: Proc. ICCHP 98 (6^{th} International Conference on Computers Helping People with Special Needs), Oldenbourg, Wien, München (1998)
4. Annamalai, A., Gopal, D., Gupta, G., Guo, H., Karshmer, A.: INSIGHT: a comprehensive system for converting Braille based mathematical documents to Latex. In Stephanidis, C., ed.: Universal Access in HCI – Inclusive Design in the Information Society. Volume 4., Mahwah, New Jersey, USA, Lea (2003) 1245 – 1249
5. Gillan, D., Barraza, P., Karshmer, A., Pazuchanics, S.: Cognitive analysis of equation reading: Application to the development of the math genie. In Klaus, J., Miesenberger, K., Zagler, W., Burger, D., eds.: Proc. ICCHP 2004 in Paris (International Conference on Computers Helping People with Special Needs). Volume 3118 of LNCS., Linz, Austria, Springer (2004) 630 – 637
6. Suzuki, M., Kanahori, T., Ohtake, N., Yamaguchi, K.: An integrated ocr software for mathematical documents and its output with accessibility. In Klaus, J., Miesenberger, K., Zagler, W., Burger, D., eds.: Proc. ICCHP 2004 in Paris (International Conference on Computers Helping People with Special Needs). Volume 3118 of LNCS., Linz, Austria, Springer (2004) 648 – 655
7. Guillon, B., Monteiro, J.L., Chécoury, C., Archambault, D., Burger, D.: Towards an integrated publishing chain for accessible multimodal documents. In Klaus, J., Miesenberger, K., Zagler, W., Burger, D., eds.: Proc. ICCHP 2004 in Paris (International Conference on Computers Helping People with Special Needs). Volume 3118 of LNCS., Linz, Austria, Springer (2004) 514–521
8. Archambault, D., Fitzpatrick, D., Gupta, G., Karshmer, A., Miesenberger, K., Pontelli, E.: Towards a universal maths conversion library. In Klaus, J., Miesenberger, K., Zagler, W., Burger, D., eds.: Proc. ICCHP 2004 in Paris (International Conference on Computers Helping People with Special Needs). Volume 3118 of LNCS., Linz, Austria, Springer (2004) 664–669

Mathematics: How and What to Speak

D. Fitzpatrick

School of Computing, Dublin City University,
Glasnevin Dublin 9
Ireland
dfitzpat@computing.dcu.ie
http://www.computing.dcu.ie/~dfitzpat

Abstract. Access to mathematical content for blind and vision impaired people continues to be a problem. The inherently visual nature of this form of presentation is neither easily or readily accessible using the linear representations in common usage by this community.

This paper proposes methodology for depicting mathematics in a non-visual manner. It will be shown how, through the prosodic component found in spoken language, the structure of mathematical formulae may be disambiguated. We will also discuss lexical cues which can be added to the utterance to further reduce the ambiguity which can be very evident in this form of material.

1 Introduction

Since its invention in the middle of the 19th century, Braille has been used to convey such diverse subjects as English, music and mathematics to blind people. Its popularity has been mainly derived from its simplicity, and its ease of production. Braille can be easily adapted to render almost anything which is available in printed form. However, there are many disadvantages to this system. Its innate simplicity does not lead to easy representation of graphical or mathematical material.

The traditional method of representing mathematical material has been to produce it in a linear fashion, akin to any other textual data. Consequently, such information has proved difficult to read.

In order to overcome some of the deficiencies in the Braille system a methodology is proposed which would harness the advances in existing technology such as synthetic speech. Nearly all of the research effort to date has been directed toward verbalising material of a "literary" nature. This, while being an essential development, still renders technical or scientific content almost totally inaccessible to those who can not see it.

The traditional reading process itself is an active-passive process, where the sighted person is the active participant while the book they are perusing is the passive partner. This process is reversed when the usual methods of auditory representation (such as audio recordings) are employed. When reading is

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 1199–1206, 2006.

performed using audio tapes or CDs the only means the reader has to become active in the reading process is to move forward or backwards in a time-line. That is, the tape can be advanced or re-wound by the "reader" but no more. Instead of being active in the reading process, the book now becomes the active participant, while the information simply flows past the passive listener. This can be very difficult when technical material is being read. It is widely accepted that the use of human readers in the provision of literary works on tape is perfectly acceptable. However, the manner in which people read scientific, or technical documents is vastly different from how novels are read.

In [3], Karshmer et al. have proven that it is possible to reverse the active-passive nature of the reading process for mathematical material. This work clearly demonstrates that using a combination of synthetic speech and an intuitive interface, the reader can take control over the information they are perusing and gain access to all components of the equation. However, this work relied on prosodically unenhanced synthetic speech to convey the material, and left the inflection of the textual representations of the equations to the synthesiser itself.

In order to devise a set of rules to prosodically enhance algebraic utterances, Stevens [9] conducted an experiment involving two experienced subjects, both of whom were native speakers of British English. Each was presented with a random set of 24 equations, and two recordings on high quality tape were made for each speaker. The participants were asked to speak the equations "...as if they were addressing a class of sighted students"[9]. They were also asked to convey the information in as neutral a manner as possible, that is, not to "... indicate any of the intentions of the mathematical notation"[9]. The recordings were analysed for pitch, timing and amplitude and these three characteristics were related to the syllabic content of the spoken versions. The results provided the designers of the system with the knowledge needed to construct a "...simplistic prosodic model"[9] which could be used to convey algebraic notation. The prosodic model used by Stevens is at odds in important ways with current views of prosodic and intonational phonology, and though it produces reasonable results, it is not the model we are using. For a discussion of appropriate phonological models and their use in synthesis, see [5,6]

In this paper, we propose a model for the verbal presentation of mathematics using prosodically enhanced synthetic speech. This model has two components; the first being the characteristics of the voice used [1], while the second is the language used to convey the material [7]. We will show, using changes to the vocal characteristic used to convey various parts of equations how mathematical content may be disambiguated. It is felt that both play an extremely important part in the disambiguation of the structure and content of equations. We realise that without a meaningful and intuitive user interface, the application of synthesis alone will not render the material intelligible or navigable. However, it is beyond the scope of this brief treatment of our work to describe our prototype user interface for the browsing of the material.

2 What Is Prosody

The prosodic component of speech is that set of features which lasts longer than a single speech sound. The term *prosody* can be traced back to ancient Greek where it was used to "refer to features of speech which were not indicated in orthography, specifically to the tone or melodic accent which characterised full words in ancient Greek" [4]. When the tonal units of ancient Greek disappeared, the use of prosody narrowed to refer to stress distinctions. By the 15th century it became known as versification; one of its primary denotations today. As the influence of the classical languages waned so also did the emphasis shift from a metric to a melodic view of prosody, though some scholars believed that the English language had no melody. However, the term *melodic prosody* remained almost forgotten until the 1940s, when it was revived as an approach to the study of linguistic analysis.

Prosody is therefore defined as "... those auditory components of an utterance which remain, once segmental as well as non-linguistic as well as paralinguistic vocal effects have been removed" or "sets of mutually defining phonological features which have an essentially variable relationship to the words selected" [4]. Therefore, it may be said that prosody contains the following aspects of speech:

– Loudness
– Duration
– Pitch
– Pausing

Acoustically, *speech* can be decomposed into three primary components; frequency, amplitude and time. "Frequency is the term used to describe the vibration of air molecules caused by a vibrating object, ... which are set in motion by an egressive flow of air during phonation." [4] The unit of measure used in the frequency domain is the Hertz (Hz). Speech is not as simple as other acoustic sounds, as it can contain many elements vibrating at different frequencies. The frequency of repetition of the complex pattern is referred to as the *fundamental frequency*, and it is this frequency which is primarily responsible for the perception of pitch. All other frequencies in any pattern are typically whole integer multiples of the fundamental frequency, and are known as the second, third, etc. harmonics.

Amplitude, (measured in decibels) is the acoustic component which gives the perception of loudness. A common definition is "the maximal displacement of a particle from its place of rest" [4]. The duration of a signal is the third component in the acoustic view. This is simply the measurement, along the time-line of the speech signal. If one considers the prosodic component of speech, then it can be reduced to a series of frequencies, a succession of intensity levels and a sequence of durations. It is these components which yield our understanding of pitch modulation, relative loudness and/or the relative duration of syllables, words or phrases.

The notion of the syllable is intrinsic to the understanding of prosody; however, it is easier to state what the syllable is not, rather than what it is. Native speakers of a language can always agree on the number of syllables any word contains, although non-native speakers can often be confused. For the purposes of this discussion, a syllable is defined (unscientifically) as a small unit into which the sounds of words are decomposed.

3 Conveying Structure

The keystone on which the presentation of syntactically complex material is based is that of conveying the structure, and grouping of an expression using pausing and alterations in the speaking rate. The paradigm on which the spoken mathematical presentation is based is that of converting a sequence of juxtaposed symbols, delimited by both white space and other visual cues (such as parentheses) into a serially transmitted linguistic approximation. In order to achieve this, a parallel was drawn between the structure found in mathematical expressions and the inherent composition of English sentences [1].

There is often a nesting and grouping to be found in language; whether spoken or written. This can be seen in the sentence, "the goal, which was scored in the last minute, won the match". The clause "which was scored in the last minute" could be omitted without a loss in intelligibility in the sentence. However, its inclusion endows a more descriptive aspect to the sentence, and conveys more information to the reader or listener than would otherwise be the case.

The nesting of clauses within sentence structures, therefore can be seen as a means to impart more descriptive semantics to the information being spoken. The subtle use of pauses and changes in prosody assist the listener in interpreting what the speaker is trying to convey. Coupled with the localised nesting of clauses within sentences, is the combination of unrelated, though logically sequenced sentences to form passages of material. Each sentence forms a unique entity, which contains both verb and noun phrases, thereby enabling its understanding. However, when placed in the context of surrounding material, the semantics of an individual sentence can be altered radically.

If we extend this to the written form of presentation, then the sequence of sentence structures can be brought together into paragraph units, which themselves contain an underlying message, of which each individual sentence forms a part. The nesting, therefore, can be seen as the overall paragraph unit, which contains sentence structures, which contains words, which themselves contain characters, both alphabetic and punctuational. By inference, therefore, it can be said that the atom of the paragraph is the character, as it is the smallest unit into which'the complex structures of English can be reduced.[1]

The function of clausal grouping was illustrated above. In a system where not only the horizontal juxtaposition of characters, delimiters and white space is permissible, the sub-grouping of component parts of the whole becomes more important. It is exactly this problem which printed mathematics presents to

[1] It is of course the case that smaller units of language exist, however, for the purposes of this discussion, and for the purposes of generating synthetic speech, the atom of the sentence is deemed to be the character.

the systems' developers, who wish to produce a verbalised approximation of the material. Unlike the methods used in written forms of natural language, mathematics uses vertical offsets, coupled with the horizontal arrangement of symbols and white space to infer different meanings to material. Consider the two examples; x^{y+z} and $x^y + z$, where exactly the same symbols are used. In the first representation, the three symbols y+z form a superscript to the x, while in the second, only the y is part of the superscript.

It is just such ambiguities that any phonological model of verbal mathematics must represent. It is beyond the scope of this brief document to outline in detail how the prosodic model caters for each individual construct found in mathematical material. It should be noted, however, that their depiction in isolation is vastly different from, and may be heavily influenced by, adjacent symbols. Accordingly, a strategy was needed to cater for the diverse combinations permissible in mathematical presentation. Unlike previous systems [8,9], the proposed model does not employ any non-speech audio to convey the material; rather it relies on the vocal representation only to impart the information. Therefore, it proved necessary to examine the grouping of the material, to deduce a method whereby it could be delivered in as unambiguous a manner as possible. Another stipulation was that the utterance be maintained at a length which would not impair the listener's ability to apprehend the material.

Previously, discussion focused on how the prosodic model is based on an application of those paradigms found in the decomposition of English; namely clauses. If one assumes, that a mathematical expression is a structure in and of itself, and that it can contain an arbitrary degree of nested sub-expressions, then it can be resolved into either a sentence structure which contains various clauses, or a paragraph (i.e., a set of sentences which may themselves be unrelated, but which when logically combined make up a superstructure). This being the case, the need was perceived to examine equations in an attempt to deduce the points at which the delimiters could be interpreted as the equivalents of clause boundaries in English. This proved to be successful, as it emerged that, in order to produce well-formatted equations, it is necessary to mark documents up using unambiguous syntactic representations. Using this as a basis, it could be inferred that, should these delimiters be present in the mark-up, it should prove possible to equate them with clause boundaries in an English verbalisation of the material.

As a consequence of this examination, it emerged that it would be possible to construct three levels of nesting.[2] Using modern speech synthesisers, it is extremely difficult to accurately control the length, and placement of pausing. Ideally, the pausing would be set at predetermined values, from which the same relative changes made to the other prosodic features of the speech could be applied to the pausing, thus maintaining compatibility between this element and the remainder of the voice contour.

[2] This can be extended. However, for the purposes of our investigation, the number of levels of indentation was kept to three, as it was consistent with the paradigm of paragraph, sentence and clause boundaries observable in English.

The two methods used to imply the grouping of terms into sub-expressions, (and by extension of sub-expressions into the whole formulae) is to insert pausing, and alter the speaking rate at strategic points within the presentation. This is not an arbitrary process but is based on both the mathematical components preceding and following the point at which the pause is needed. For example, the expression $a + b + c$ is a simple, linear concatenation of three quantities, separated by two relational operators, irrespective of whether the material is being read visually or auditorily. However, the expression $\frac{a}{b} + c$ is non-linear in the print medium, but linear in the audio domain.

Accordingly, something must be introduced to indicate to the listener that the fractional component of the expression merely encompasses the first two elements, and that the final one is not contained therein. The method whereby this is achieved, is to speak the expression as follows: "a over b, plus c". Using this method, it can be clearly seen that the distinction between the fractional component of the expression and the remaining relational operator and quantity is evidenced. If one also adds a slight quickening of the speaking rate of the utterance of the fraction, then the distinction becomes even more apparent to the listener.

4 Complex Formulae

The example of the previous Section is an extremely simple demonstration of the capabilities of the methods used in the system to produce spoken output from mathematical material. This method works equally well when complex entities are included in the expressions. As an illustration of a more complex example consider the following equation:

$$\sum_{i=1}^{n-1} a^i + \frac{i+1}{i-1} \tag{1}$$

Here it can be seen that this expression contains several operators in combination to produce a formula which could prove quite difficult to read. Using unadorned text strings, it would not be discernible where the scope of the various superscripts ended, or which items were contained in the fractional component of the formula. Also, were alterations in the prosody alone to be used, there could be confusion as to whether the summation encompassed the entire scope of the remaining elements or whether it was merely confined to that component which immediately followed it. Using the prosodic enhancements, in combination with lexical cues, it is possible to produce an audio rendering which can clearly and intelligibly inform the user of the means whereby this formula can be decomposed.

The first aspect in producing any meaningful output from mathematical material is to determine the degree of nesting in the formula. Once this has been established, the levels, and indeed lengths of pausing required to speak the content can be determined. In the preceding example, the summation encompasses the entire gamut of the remaining material, thereby yielding one level of nesting. The material itself consists of one superscripted element, followed by a fractional component. However, the fractional component itself comprises two complex terms, as the numerator consists of $i + 1$, while the denominator consists of $i - 1$. Accordingly, there are several degrees of nesting observable in this expression. There must be a lengthy pause to indicate that the entire formula is contained within the scope of the summation, followed by various pauses to indicate to the listener to the scope of the superscripts, and sub-expressions of the fractional element.

The approximate verbal rendering of this equation would be as follows. For simplicity, the symbol ".." has been used to indicate a longer pause, more akin to that found between paragraphs.

> "sum from i equals 1 to n minus 1 of.. a to the i plus 1 plus. begin Fraction, Begin numerator i plus1 end numerator over, begin denominator, i minus 1, end denominator end fraction."

The pausing in this example illustrates an approximation of the nesting used to indicate the scope of various operators. Combined with increases, and decreases in the speaking rate between the various clause boundaries, it has proved possible to use this mode of presentation to determine the scope of operators in expressions even as complex as this. The lexical cues used in this example are based on those defined in [7], and while they lengthen the utterance somewhat they further reduce the ambiguities in such an expression as this. Moreover, the use of pausing *after* operators is a useful aid in assimilating the fact that the operator is itself followed by a complex term.

In the quoted passage above, the reader can observe the pause after the lexical operator "over". Once the listener has become familiar with the system, this pausing strategy can be used to anticipate the fact that a complex term is about to be spoken. Conversely, in a simple equation, the pausing would not be placed after the operator but before. This indicates the scope of operators (such as fractions), and also informs the user, without the addition of extra lexical information, that the term they are about to hear is a simple one.

The alterations in the speaking rate have been carefully calibrated to ensure that, even at the deepest level of nesting, the utterance is not delivered too rapidly. With this in mind, a value of 6% has been chosen to assist in informing the listener of the subtleties of the grouping which are so easily observable to the visual reader. The reason that an increase of 10% were not chosen, is that if three or more levels of nesting were encountered in an expression, then the ultimate speaking rate would be too fast, and as a consequence, the listener would be unable to ascertain what was being spoken.

5 Conclusion

Though the inclusion of prosodic alterations in the utterance of mathematical equations alleviates the problem of apprehending the material, it is not a solution of itself. Just as the written language of printed mathematics is a form of presentation, so also is prosodically enhanced synthetic speech. It is *imperative* that an intuitive interface be placed on top of this presentational modality [3,9,1] to facilitate rapid and easy perusal of the material. We also believe that further enhancements can be made to the actual "spoken language of mathematics". Such alterations would include devising alternative lexical cues to attempt to create a more abbreviated utterance.

References

1. D. Fitzpatrick. *Towards Accessible Technical Documents: Production of Speech and Braille Output from Formatted Documents*. PhD thesis, School of Computer Applications, Dublin City University, 1999.
2. A. I. Karshmer and D. Gillan. *How well can we Read Equations to Blind Mathematics Students: Some Answers from Psychology* Proceedings of the 2003 Human Computer Interface International Conference, Crete, Greece, July, 2003.
3. A. I. Karshmer, C. Bledsoe, and P. Stanley. *The Architecture of Comprehensive Equation Browser for the Print Impaired* vlhcc, pp. 319-320, 2005 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC'05), 2005.
4. E. Couper-Kuhlen. *An Introduction To English Prosody*. Edward Arnold Ltd., 1986.
5. D. R. Ladd. *Intonational Phonology.* Cambridge: CUP, 1996.
6. A. Monaghan. *Intonation in a Text-to-Speech Conversion System.* PhD thesis, University of Edinburgh, 1991.
7. A. Nemeth. *Abraham Nemeth's Anthology on the Unified Braille Code* Available at `http://www.math.virginia.edu/ arm4r/nemeth`
8. T. V. Raman. *Audio Systems for Technical Reading*. PhD thesis, Department of Computer Science, Cornell University, NY, USA, May 1994.
9. R. Stevens. *Principles for the Design of Auditory Interfaces to Present Complex Information to Blind People*. PhD thesis, Department of Computer Science, January 1996.

SBT: A Translator from Spanish Mathematical Braille to MathML

Fernando Alonso, José L. Fuertes, Ángel L. González, and Loïc A. Martínez

Facultad de Informática, Universidad Politécnica de Madrid, Campus de Montegancedo, 28660-Boadilla del Monte, Madrid, Spain
{falonso, jfuertes, agonzalez, loic}@fi.upm.es

Abstract. One of the key issues for integrating blind people into everyday life is education. For many educational subjects, however, there is no suitable assistive technology for blind people. One especially sensitive question facing blind people is learning mathematical language, apart from interacting with sighted teachers and students. This article presents a Spanish mathematical Braille to mathematical notation (MathML) conversion system to fill this gap. The translation system has been designed as a portable programming library, it solves ambiguities in Spanish mathematical Braille and it generates an intermediate code, which means that it can be easily adapted to other Braille languages or to other output formats.

1 Introduction

The integration of disabled people and, especially, blind people into working life in society is now a challenge into which international institutions and states are pouring a lot of resources [1,2,3]. But for this integration to be effective, it needs to start at school and blind children should receive the same education as sighted people.

There is now a range of assistive technology for blind or visually impaired people (Braille displays, screen magnifiers, screen readers [4,5], etc.), which is a big help for their social integration and training processes. However, there are many areas, which are under research, in which blind people still do not have suitable assistive technology for this integration, including, for example, urban mobility around streets, buildings and on public transport, or equal access to all mobile telephony services.

One especially sensitive issue concerning the education of blind learners is mathematical language learning and interaction with sighted teachers and students.

At present, the best way for a blind person to read and write mathematical signs is using the mathematical Braille language, but sighted people are unable to understand this notation. On the other hand, a blind person cannot use normal (typeset) mathematical language for obvious reasons. All this erects a communication barrier between blind learners and their sighted teachers and classmates, which has a negative impact on their training.

This article aims to solve this problem for the Spanish-speaking community by developing a Spanish mathematical Braille [6] to MathML [7] translator. This translator is part of a platform that will be able to convert several mathematical Braille notations into several computational mathematics notations and vice versa.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 1207 – 1214, 2006.

As discussed in the Related Work section, a number of projects on processing mathematical Braille in different languages have got under way over the last twenty years, but none of them have dealt with the Spanish notation.

Section 2 of the paper gives an overview of related work in the field of translation between mathematical Braille and other languages. Section 3 gives a brief introduction to Spanish mathematical Braille notation, also discussing the ambiguities arising when it is computer processed. Section 4 introduces the Symbolic Braille Translation (SBT) system, which is the focus of this article and converts mathematical Braille into MathML. Finally, section 5 states the conclusions.

2 Related Work

Several R&D projects have been developed or are under development for languages other than Spanish. The work can be classified according to the "direction" of the translation. So, we have:

- Mathematical notation to mathematical Braille translators: This group includes LaBraDoor [8], which translates LaTeX to Marburg Braille Notation (used in Germany); MAVIS [9,10], which is a formula browser that allows formula navigation using voice output, and math2braille [11], which is an opensource architecture for making translators from MathML to Braille.
- Mathematical Braille to mathematical notation translators: These include Insight [12], which converts formulas from Nemeth Braille code (used in the United States) to LaTeX, and MMBT (Multi-Language Mathematical Braille Translator) [13], which is an ongoing opensource project that will allow translation from French (historical and new notation), British and Italian Braille notations to LaTeX and MathML.
- Two-way translators: Universal Maths Conversion Library [14], which is an ongoing project aiming to create a library to support the creation of mathematical notation to mathematical Braille translators and vice versa. The supported Braille notations will be Nemeth, Marburg, French (historical and new notation), British and Italian Braille.

There are other projects designed to make mathematics accessible for blind people. These include the LAMBDA project [15], whose goal is to define a linearised mathematical code (Lambda Math Code) derived from MathML that can be used with assistive technologies for blind people.

It is clear from the above that little effort has been invested in the Spanish notation, and a Spanish mathematical Braille translator was a growing need for the Spanish-speaking blind community.

3 Spanish Mathematical Braille and Its Ambiguities

By 1987, the Spanish Braille system used for texts composed of alphanumerical and punctuation characters was relatively standardised, but there were a wide variety of different notations in fields such as mathematics, because they were generated and used as needed. In June 1987, representatives of the Braille presses met at Montevideo (Uruguay) with the aim of reaching agreement on the use and unification of the Spanish Braille notation, mainly with respect to mathematical symbols [6]. The result was the standardisation of the Spanish Braille system for both basic and scientific notations.

Blind people can use and have no problems interpreting the standardised mathematical Braille, because they can settle any doubts they may have from the context of the information that they are reading. However, the actual Braille notation does raise serious problems of ambiguity when it is computer processed as ASCII-Braille code (in which each Braille character can be matched to an ASCII character).

There are a number of ambiguities that occur when trying to convert a mathematical ASCII-Braille text to print and which need to be removed. The ambiguities are:

- *Text and formulas:* The ambiguity arises when trying to make a distinction in ASCII-Braille between the two types of information (text and formulas or mathematical symbols) in unified code. This distinction is crucial, because a character or character set can have a completely different meaning depending on whether it appears in a text or mathematical formula. But mathematical Braille does not specify how to differentiate the transition from one to another.
- *Ordinal numbers:* The feminine gender of the Spanish ordinal numbers is formed by the character "a". Ambiguity arises in this case because it is impossible to tell whether you are processing an ordinal number or a fraction. For example, if processing the ASCII-Braille text "`#,>a`", there would be two possible printed representations shown in formula 1.

$$\frac{10}{1} \text{ or } 10^{a} . \tag{1}$$

- *Enclosure signs:* The ambiguities that arise in ASCII-Braille with respect to these signs are indicated in Table 1.

Table 1. Enclosure signs

<table>
<tr><th>ASCII-Braille</th><td colspan="2">2</td><td colspan="2">(</td><td colspan="2">)</td><td colspan="2">@l</td><td colspan="2">@|</td><td colspan="2">2,</td></tr>
<tr><th>Possible characters</th><td>(</td><td>ê</td><td>[</td><td>á</td><td>]</td><td>ú</td><td>{</td><td>|</td><td>⎨</td><td>∍</td><td>⎬</td><td>∈</td></tr>
</table>

- *Exclamation marks and speech marks:* The open exclamation mark[1] "¡", the close exclamation mark "!" and the add operator are all represented by the ASCII-Braille symbol "+". Speech marks """ and the multiplied by operation are represented by the symbol "<".

All these ambiguities need to be removed to automatically convert ASCII-Braille into print. The next section, describing the translator from ASCII-Braille to the mathematical language MathML, discusses how these ambiguities are resolved.

---

[1] In Spanish, an upside down exclamation mark is used to indicate where the exclamation begins in a similar way to how speech marks are used in English.

4 Mathematical Braille to MathML Translator

The strategy followed to build the translator was based on compiler theory, using a general-purpose design, which would allow rapid reuse of most of the translation system. The input for the analysis module is mathematical ASCII-Braille, which is transformed into intermediate code (in the shape of triples) and is then translated to MathML. This is a highly reusable schema, as it is easy to change the front-end of the translator to admit other Braille languages or the back-end to generate other mathematical notations.

The main steps taken to solve the problem are described in the following.

4.1 Resolving Ambiguities

After analysing the different types of identified ambiguities explained in section 3, it was decided in conjunction with the Spanish National Organisation for the Blind (ONCE) that they should be removed by using the Braille dot 7 for the least common symbols[2]. Accordingly, the representation for the more frequently used option remains unchanged. To make the system more agile, the decision was taken not to opt for interactive translation, which would involve asking the user.

Following this philosophy, the ambiguities were resolved as follows:

- *Text and formulas*: To distinguish text from mathematics, we decided to create some special codes to indicate the start and end of the mathematical symbology. These codes are: "@@xy" and "@@&x" (in ASCII-Braille).
- *Ordinal numbers*: To distinguish feminine ordinal numbers from fractions, we decided to use the ASCII-Braille character "A" (that is, the character "a" plus dot 7) to refer to the ordinal. This way we have:
  - "#,>a" will be $\frac{10}{1}$
  - "#,>A" will be 10^{a}.
- *Enclosure signs*: We decided to use the character plus dot 7 to represent "(", "[", "]", "{", "⎨" and "⎬", to give the ASCII-Braille characters "Ê", "[", "]", "@L", "@Ä" and "2Å", respectively.
- *Exclamation marks and speech marks*: These ambiguities in the textual part have been resolved using dot 7 to represent exclamation and speech marks, giving the ASCII-Braille characters "!" and "Ì".

Thanks to this way of settling ambiguities, users can quickly and easily interpret this new Braille notation.

4.2 Designing the Intermediate Code

The most noteworthy module in the translator design is the intermediate code generator, sandwiched between the analysis phase and code generator (Fig. 1). Its input is

---

[2] Braille is composed of 6 dots placed in two columns of 3 dots each. Depending on whether or not the dots are raised, they represent one character or another. Another two dots (called dots 7 and 8) are added in computerised Braille and one is placed under each column, a Braille character now being composed of two columns of four dots each.

aparser tree and it outputs an intermediate code (which we have designed as triples) independent of the source and object languages. The triples can generally and independently represent the elementary operations and their operands. The advantage of using the triples is that these source and object languages can be changed without modifying all the other modules of the translator.

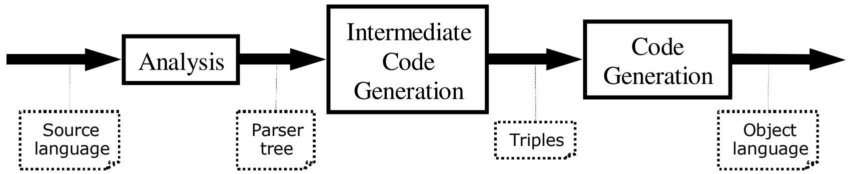


Fig. 1. Translator design

4.3 Implementing the Translator as a Library

It was decided to implement the translator as a library that could be called up from different interfaces or applications.

Owing to the diversity of platforms on which this library might be needed, the requirement of portability to any platform equipped with a compiler conforming to the ANSI C++ standard was established for construction.

The result is a C++ library that provides a translation of a file in Spanish mathematical ASCII-Braille to MathML. This library has been developed to work in the same way on DOS, Windows and Linux.

4.4 Developing the User Interfaces

Two applications were built around the translator library. The first is a compact command line-based application for the DOS and Linux operating systems. The second is a system that provides a window-based interface and works in the Windows environment.

Command-line SBT. Its purpose is to provide a compact tool with low memory and disk storage requirements to conform to the requirements of small portable Braille systems.

The operation involves calling the SBT by indicating the file to be translated and the output file. The system issues a series of messages to tell the user what is going on at any time. The user is also informed about any error detected during the translation process, to determine what problems there are with the input file. Personal computers running DOS or Linux will need a text-mode screen reader to allow this information to reach the user via voice synthesiser or Braille displays.

SBT for Windows. Its purpose is to provide an integral prototype for both editing and visually representing symbolic Braille. It was developed in compliance with the Spanish standard on software accessibility UNE 139802:2003 [16], so the blind user can operate it by using a screen reader.

There is a menu to perform a range of tasks: loading a file, translating and visualising the results, and inserting text or formulas in ASCII-Braille. There are two areas underneath the menu (Fig. 2). An editing area (top half of the screen) visualises the

ASCII-Braille content of the open file that can be modified by the user. A visualisation area (bottom half of the screen) shows the translation results. This area displays the information based just on the MathML file obtained by the translator.

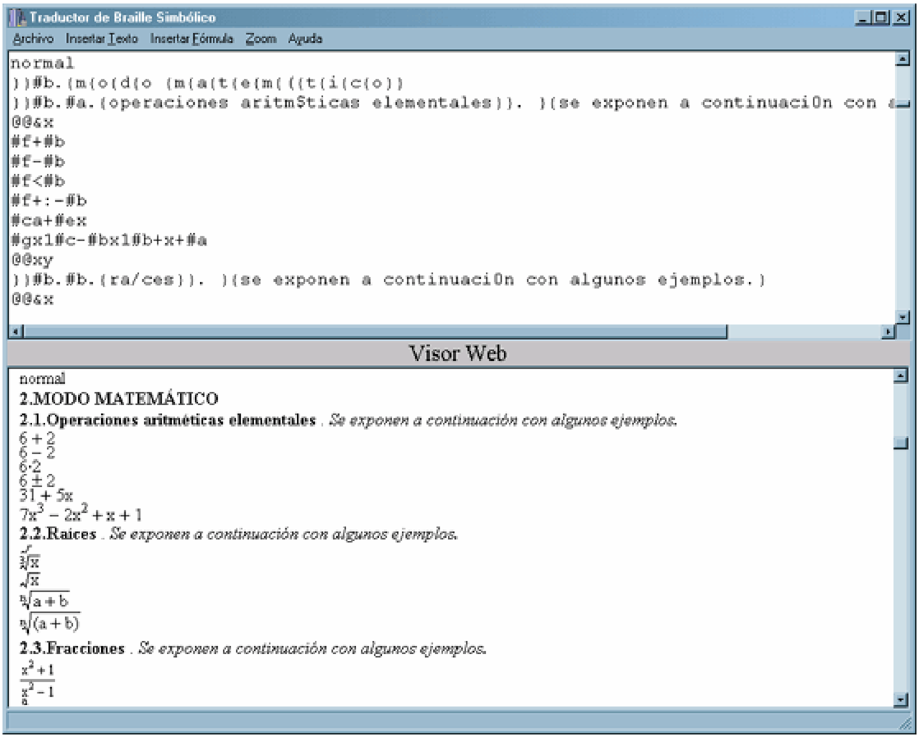


Fig. 2. Windows version of the SBT

5 Conclusions

The mathematical Braille to mathematical language translator system described in this paper meets the Spanish-speaking community's need for mathematical Braille to MathML translation, and can generate XML documents that retain the semantics in which the mathematical expressions are embedded. This opens up a new communication channel between sighted and blind people, who can participate fully in the mathematical teaching/learning process.

Additionally, it has the added value, thanks to the translator module design philosophy (Fig. 1), of being able to separate the mathematical Braille language variant (the source code) from the mathematical language generated by the translator (the object code). As a result of the definition of the intermediate representation as triples, the creation of other translators using any Braille language (like English Braille, Nemeth...) to MathML would involve building the analysis and intermediate code generation phases, as shown in Fig. 3. Furthermore, if instead of generating MathML,

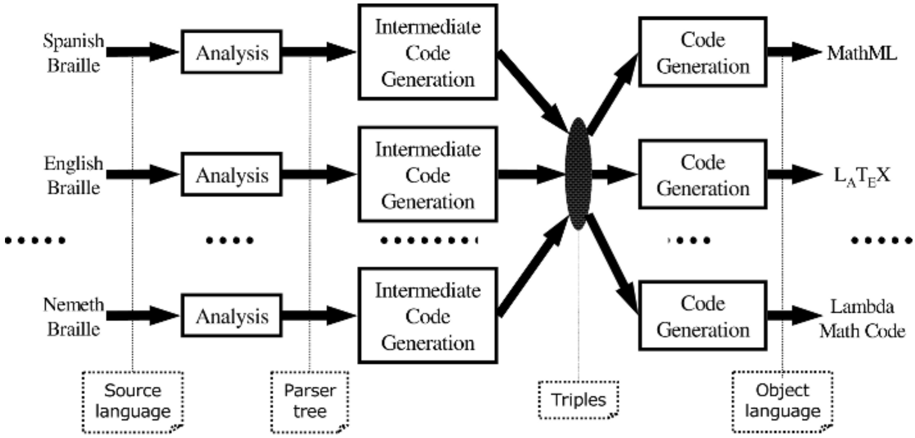


Fig. 3. Overall SBT architecture

we wanted to translate the Braille language to other mathematical notation languages (like $\mathrm{L_AT_EX}$, Lambda Math Code...), all we would have to do is build the respective code generator to read and translate the triples to the selected code.

References

1 Ley de igualdad de oportunidades, no discriminación y accesibilidad universal de las personas con discapacidad (Law on the equality of opportunities, non discrimination and universal access for people with disabilities). Boletín Oficial del Estado. Spain (2003)

2 International classification of functioning, disability and health (ICF), Nonserial Publication. World Health Organization (2001)

3 European Year of People with Disabilities 2003. Official Journal of the European Communities, Vol. 335. European Union (2001) 15-20

4 JAWS 7.0: Quick Start Guide. Freedom Scientific. Florida (2006). Available: `http://www.freedomscientific.com/fs_support/documentation/JAWS7_0Docs/JAWS70.txt`

5 Baum Engineering SRL. Welcome to Gnopernicus. Timisoara, Romania (2005). Available: `http://www.baum.ro/gnopernicus.html`

6 Della, J. J., Ortiz, M. A., Domínguez, F. R., Roig, C.: Código matemático unificado para la lengua castellana. Aprobado por la reunión de imprentas Braille de habla hispana. Montevideo (1987)

7 Mathematical Markup Language (MathML) version 2.0. W3C Recommendation. 2nd ed. (2003). Available: `http://w3.org/TR/2003/REC-MathML2-20031021/`

8 Batusic, M., Miesenberger, K., Stoeger, B.: Labradoor, a contribution to making mathematics accessible for the blind. In: Proc. ICCHP 98. Oldenbourg (1998)

9 Karshmer, A., Gupta, G., Geiger, S., Weaver, C.: Reading and writing mathematics: The mavis project. Behavior and information technology 18 (1998) 2–10

10 Karshmer, A., Gillan, D.: How well can we read equations to blind mathematics students: some answers from psychology. In: Stephanidis, C. (ed.): Universal Access in HCI– Inclusive Design in the Information Society, Vol. 4. Mahwah, NJ (2003) 1290–1294

11 Crombie, D., Leonor, R., McKenzie, N., Barker, A.: math2braille: Opening Access to Mathematics. In: Proc. ICCHP 2004 (International Conference on Computers Helping People with Special Needs). Lecture Notes in Computer Science, Vol. 3118. Paris, France (2004) 670-677
12 Annamalai, A., Gopal, D., Gupta, G., Guo, H., Karshmer, A.: INSIGHT: a comprehensive system for converting Braille based mathematical documents to $\LaTeX$. In: Stephanidis, C. (ed.): Universal Access in HCI– Inclusive Design in the Information Society, Vol. 4. Mahwah, NJ (2003) 1245–1249
13 Moço, V., Archambault, D.: Automatic translator for mathematical Braille. In: Stephanidis, C. (ed.): Universal Access in HCI– Inclusive Design in the Information Society, Vol. 4. Mahwah, NJ (2003) 1335-1339
14 Archambault, D., Fitzpatrick, D., Gupta, G., Karshmer, A. I., Miesenberger, K., Pontelli, E.: Towards a universal maths conversion library. In: Proc. ICCHP 2004 (International Conference on Computers Helping People with Special Needs). Lecture Notes in Computer Science, Vol. 3118. Paris, France (2004) 664-669
15 Nicotra, G.: LAMBDA Project (2006). Available: `http://www.lambdaproject.org`
16 Health informatics. Computer applications for persons with disabilities. Computer accessibility requirements. Software, UNE Standard 139802:2003 (2003)

Braille Math Made Easy with the Tiger Formatter

John A. Gardner[1], Leon Ungier[1], and John J. Boyer[2]

[1] ViewPlus Technologies, Inc.
1853 SW Airport Ave.
Corvallis OR 97333, USA
{john.gardner, leon.ungier}@viewplus.com
[2] JJB Software, Inc
913 Acewood Boulevard #109
Madison, WI 53714, USA
john.boyer@jjb-software.com

Abstract. The Tiger® Braille Formatter is a Windows® software application used to transform MS Office documents into a form suitable for embossing on ViewPlus® Tiger® technology embossers. Tiger® embossers have 20 dot per inch resolution and can emboss Braille and tactile graphics with variable height dots. The Formatter's primary purpose is to replace text with Braille and reformat the document to accommodate the considerably larger font size needed for Braille than normal text. A new release of the Formatter will include ability to transform math within MS Word documents to virtually any standard Braille code, to DotsPlus® Braille, or to custom Braille codes developed by individual users. The new open source Liblouis open source translator is used for both literary and math Braille translation.

1 Introduction

The Tiger® Braille Formatter software application was developed so that users of ViewPlus® embossers could easily convert text in MS Office documents to Braille before embossing. Most MS Word and Excel documents can be translated to Braille and reformatted automatically with no special editing by the user. Formatter permits many user options including language. Presently more than 20 literary Braille languages are supported, and the list is growing. Formatter is currently being expanded to provide translation of math, in the form of MathML, to various Braille math codes as well as to DotsPlus Braille. The Formatter application, DotsPlus® Braille, and the open source liblouis Braille translation engine used by Formatter are described here.

2 Tiger® Braille Formatter

Tiger® Braille Formatter is a plug-in for the Microsoft Office® Word and Excel applications. Word or Excel users may select a broad variety of Braille translation options, including choice of Braille language, choice of non-contracted or whatever levels of contracted Braille are permitted by the selected language, choice of using

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 1215 – 1222, 2006.

capital letters, emphasis markup, etc. Users may choose to see dots or ASCII font characters on screen after formatting is done and may choose to see the original text displayed as interline regular text. The transformed file may be embossed on any ViewPlus® embosser. The regular text can also be printed on the embossed copy if the ViewPlus® Pro Ink or Emprint™ embosser is used.

Some special Braille formatting needs are provided, e.g. ability to show page numbers of both the final Braille document and the original print document. However the overall formatting of margins, paragraph breaks, footer and header type, etc. is determined by the Office® application and is not altered by the Braille transformation process. Consequently, easily-readable Braille documents can be prepared by anybody who knows how to use standard computer software. Expert Braille transcribers may hand-correct the many small details required to conform to all the requirements of official Braille codes, particularly those that depend on context that cannot be recognized by computers.

Graphical information is retained and embossed. By default, dark areas of graphics are embossed with high dots, and light areas with low dots. Text within graphics is translated to Braille if the text is represented by proper screen fonts. Bit mapped text in graphics is not recognized as text and is not translated. In order to represent such text in Braille, a sighted person can replace it with proper text in a standard text box placed to cover the original bit map image. The text in this text box will then be properly translated.

Human editing may be needed in some other cases as well, e.g. reformatting of tables that are too large to fit when text is translated to Braille; Braille text on graphics that is too large to fit in the original position.

3 DotsPlus® Braille in Math Equations

DotsPlus® Braille [1,2] is an extension of standard Braille that represents letters, numbers, and a small number of other characters as Braille symbols and represents most other characters as graphic images. Most non-alphanumeric symbols such as the plus, slash, and backslash are tactile images of the print symbol. Punctuation marks are graphic symbols similar either to the print or Braille symbol. Accented letters of non-English languages are represented by graphic symbols that feel like the Braille patterns in those languages.

Lower case Roman letters are represented by the standard internationally-recognized Braille patterns. Upper case Roman letters can be represented by double cells consisting of the lower case Braille cell preceded by either a dot-6 (English standard) or dot-46 (used as the capital letter prefix in most non-English languages). These are called regular DotsPlus® and DotsPlus® Euro respectively. Capital letters may also be represented as single 8-dot cells in which the upper 6 dots are the lower case Braille cell with an additional dot on the bottom row left (dot-7 position). This is DotsPlus® Expert. Greek letters are also representable as double cells or single 8-dot cells with a dot-8 (dot on lower right). DotsPlus® numbers are those of Central European computer Braille for which 1-9 are the letters a-I with an additional dot-6. Zero is dot-346.

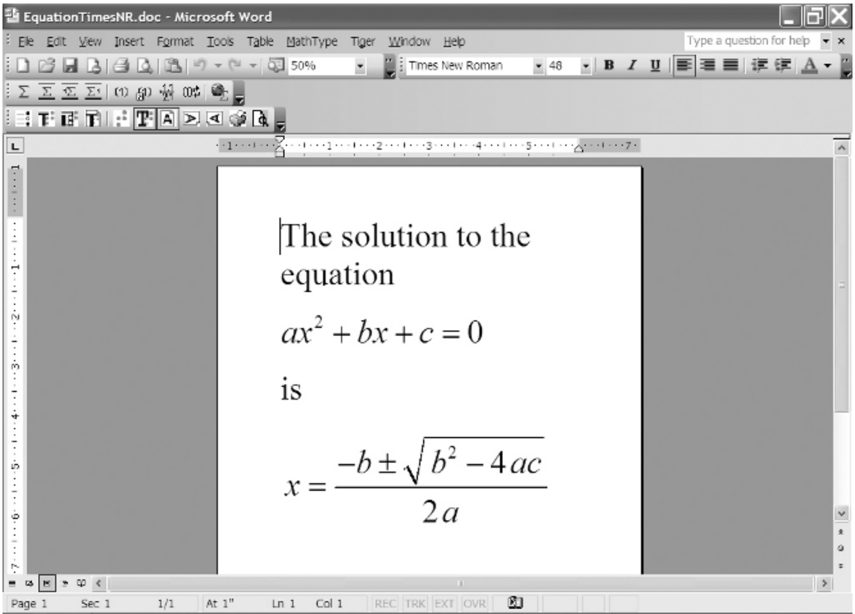


Fig. 1. Screen image of a MS Word file showing the solution to the quadratic equation

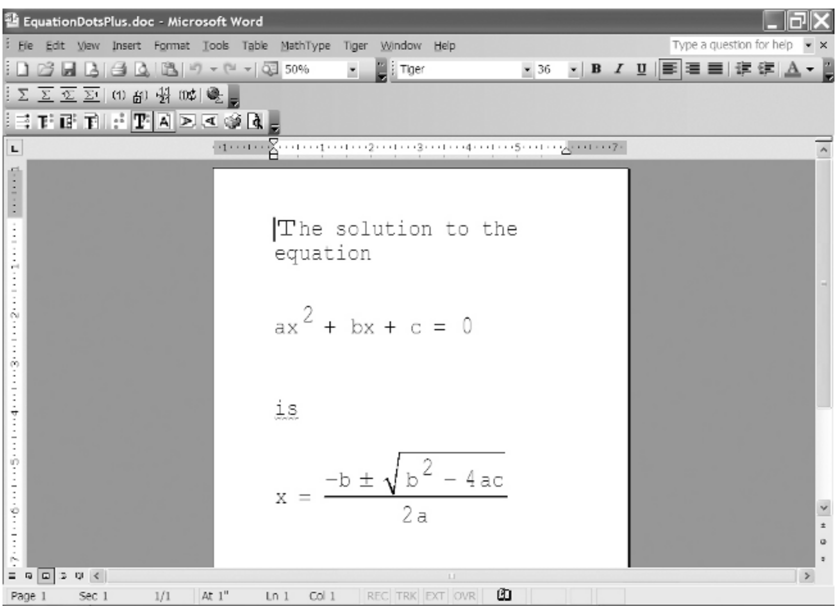


Fig. 2. Screen image of the file of Figure 1 shows that page when text has been converted to the Tiger® screen font which prints as DotsPlus® Braille and the equations have been displayed with the MathType® Tiger® environment

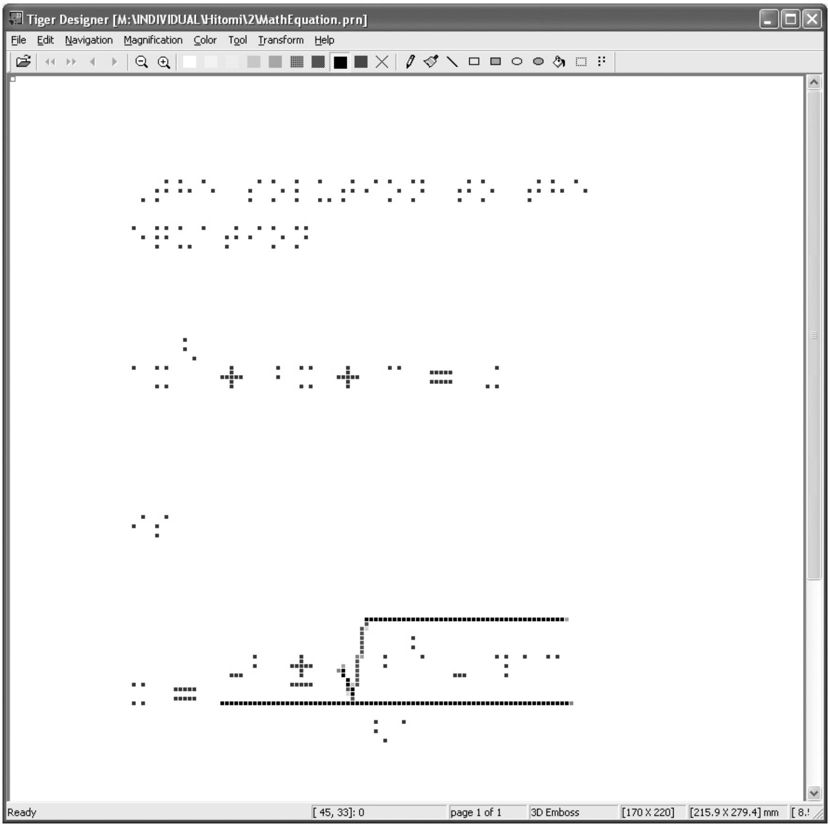


Fig. 3. The pattern of embossed dots when Fig. 2 is printed to a ViewPlus® embosser. The regular literary and math text and the math graphic portions could also be printed in ink on the embossed image with a Pro Ink or Emprint™ embosser.

DotsPlus® Braille can be used for anything, but it was originally developed to permit math to be represented tactually in the two-dimensional form normally used by sighted people. This form is not as spatially efficient as math Braille. The two-dimensional standard math layout is also less efficient for sighted people than a linear form. The less efficient 2-D form is used because it is intuitively more understandable than a linear form. There is no reason to believe that the same is not true for blind readers.

It is quite easy in principle for mainstream math to be embossed in "normal" form when DotsPlus® Braille is used for characters in the equation. The math display engine needs to have enough flexibility to permit screen fonts to be changed to a font that embosses as DotsPlus®. The very popular MathType® math editor used in MS Word and many other Windows® applications permits equations to be displayed in such a form. Version 2.2 of the MathPlayer plug-in for Internet Explorer also permits equations to be formatted to emboss as DotsPlus® [3]. One selects a Tiger®, Tiger® Euro, or Tiger® Expert environment, and the equations are transformed to ones that emboss as regular DotsPlus®, DotsPlus® Euro,or DotsPlus® Expert respectively. Figure 1 shows a math example page. Figure 2 shows that page when text has been converted to the Tiger® screen font which prints as DotsPlus® Braille and the equations have been displayed with the MathType® Tiger® environment. Figure 3 shows the embossed printout when the file shown in Figure 2 is printed to a ViewPlus® embosser.

DotsPlus® math representation has been used successfully by a handful of blind university students[4-6] in the US and UK for a variety of reasons. Some were unable to learn to read math in the conventional math code, and others used DotsPlus® because it could be obtained almost immediately from MS Word documents at a time when translation to standard math code was tedious and time-consuming. Others use it because it more easily facilitates communication with sighted people. The standard keyboard characters of DotsPlus® Braille can be learned in minutes by a proficient Braille reader. More advanced math symbols are learned when needed, just as they are by sighted readers. A blind student who has no comprehension of the way standard math is laid out will need to learn math notation just as sighted students do. The students who have learned and used DotsPlus® Braille have a wide range of ability but none has experienced a serious learning curve when learning to read math with DotsPlus®. The same certainly cannot be said for standard math Braille code learners.

4 Official Math Braille

Most major countries have adopted some official representation for math in Braille. Generally these math Braille codes differ markedly from country to country, even those sharing a common language. The US Nemeth math Braille code, for example, is very different from the UK Braille math code. The only commercially-available software that transforms a mainstream math format to standard math Braille is the Duxbury DBT Braille translator (`http://www.duxburysystems.com`). Math written in Scientific Notebook (`http://www.mackichan.com/`) can save files as LaTeX which can be translated to US, UK, or French math Braille by DBT. A number of other math Braille translation projects have been reported [7-11] but none have yet resulted in applications that are presently available in a form that is useful to most end users.

5 The Liblouis Open Source Braille Translator Project

The Liblouis Braille project was begun as a commercial collaboration with the intention of developing a multi-language open source Braille translator. In general there are two facets to Braille rules – content and formatting. Achieving excellent Braille content is considered easier than achieving excellent formatting, so the initial goal of Liblouis has been to concentrate on obtaining excellent translation of text content to Braille in multiple languages. This capability is adequate for the Tiger Braille Formatter, since format is done by settings of the Office document. It also serves the needs of a broad group of other companies who make Braille devices of many kinds. It is hoped that this community of interest will broaden to include major Braille agencies.

These agencies are best equipped to assure that the translation tables for their particular language are maintained as rules change. Liblouis information is currently available through the JJB web site `http://www.jjb-software.com`. It will eventually be moved to an appropriate open source host, and pointers to the open source site will be available from the above JJB site.

A new library, Liblouisxml, has recently been added to the Liblouis family to translate from MathML to various math Braille codes. Nemeth and UK math code is the initial target, and it is expected that tables for European math Braille codes will be developed through collaboration with experts on those Braille codes.

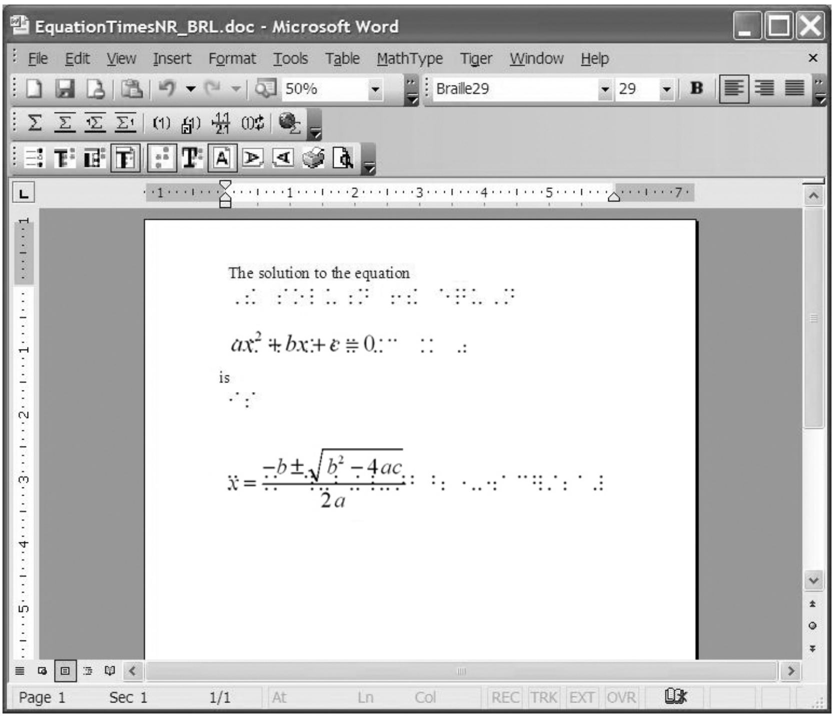


Fig. 4. Screen image of the file of Figure 1 after formatting for a Tiger® embosser. The Braille options are: Contracted US literary Braille and Nemeth math Braille. An option has also been selected for showing original text. When embossed on the ViewPlus® Pro Ink or Emprint™ embosser, the standard text is printed in ink, and the Braille dots are embossed. When printed to another ViewPlus® embosser, dots are embossed but the standard text is ignored.

Figure 4 shows a screen shot of the translation of the file of Figure 1 when the US Nemeth Braille code is selected for math. The figure shows both the Braille dots and regular text. When embossed on the ViewPlus® Pro Ink or Emprint™, the regular text can be printed in ink. The Braille dots will be embossed by those two or any other ViewPlus® embosser.

6 Testing

One of the authors (JJB) is an expert on US Braille and has served as primary alpha tester for the Braille output of Liblouis. A number of blind ViewPlus® users, consultants, and ViewPlus® distributors serve as beta testers. Braille errors reported by these beta testers are continually corrected by improving the translation tables for that language. Some errors in non-English Braille required addition of commands to the basic Liblouis translator routine. Some languages have been tested more thoroughly than others. As of this writing, the authors believe that Liblouis accuracy is excellent for US and UK English and for several Scandinavian languages. Most standard math is being translated properly, but very arcane advanced math expressions still remain to be tested thoroughly.

Acknowledgements

The authors thank Mr. Thomas Johnston for assistance with creating UK and non-English tables and for valuable assistance with beta testing.

References

1. Gardner, J., (2002) Access by Blind Students and Professionals to Mainstream Math and Science, Proceedings of the 2002 International Conference on Computers Helping People with Special Needs, Linz, Austria, July 15-20, 2002
2. Gardner, J. (2003) DOTSPLUS Braille Tutorial, Simplifying Communication Between Sighted And Blind People, Proceedings of the 2003 CSUN International Conference on Technology and Persons with Disabilities, Los Angeles, CA, March 19-22, 2003
3. Gardner, J., Soiffer, N., and Suzuki, M. Emerging Computer Technologies For Accessible Math, Proceedings of the 2006 CSUN International Conference on Technology and Persons with Disabilities, Los Angeles, March 21-25, 2006
4. Francioni, J., and Smith, A., (2002) Computer Science Accessibility for Students with Visual Disabilities, Paper published in proceedings of *33rd SIGCSE Technical Symposium on Computer Science Education*, Northern Kentucky, February 2002, pp. 91-95
5. Francioni, J., (2005) Observations on the Math Needs of College Undergraduates with Visual Disabilities, Proceedings of the National Federation of the Blind GAMA (Goals for Achieving Math Accessibility) Summit, April 14-15, `http://www.nfb.org/nfbji/goalsmath/mathneeds.htm`
6. Gardner, J. (2005) New Technologies for Accessible Tactile Math and Accessible Graphics, Proceedings of the National Federation of the Blind GAMA (Goals for Achieving Math Accessibility) Summit, April 14-15, 2005 `http://www.viewplus.com/products/braille-math/tactile-spatial-math/`
7. Batusic, M., Miesenberger, K., Stoeger, B.: Labradoor, a contribution to making mathematics accessible for the blind. In Edwards, A., Arato, A., Zagler, W., eds.: Proc. ICCHP 98 (6th International Conference on Computers Helping People with Special Needs), Oldenbourg, Wien, München (1998)
8. Mo¸co, V., Archambault, D.: Automatic translator for mathematical Braille. In Stephanidis,C.,ed.:Universal Accessing HCI –Inclusive Design in the Information Society. Volume 4.,Mahwah,NewJersey,USA,Lea(2003) 1335 – 1339

9. Moço V. and Archambault D.: A transcription tool for mathematical Braille. In Craddock G., McCormack L., Reilly R., Knops H., Eds.: Proc. of the AAATE'03 Conference in Dublin, Ireland. - IOS Press, Amsterdam, The Netherlands (2003) 481 - 485
10. Moço V., and Archambault, D. (2004) Automatic Conversions of Mathematical Braille: a Survey of Main Difficulties in Different Languages, Proceedings of the ICCHP Conference, Paris, France, July 7-9, 2004
11. Archambault, D., Fitzpatrick, D., Gupta, G.,Karshmer, A., Miesenberger, K. and Pontelli P. (2004) Towards a Universal Math Conversion Library, Proceedings of the ICCHP Conference, Paris, France, July 7-9, 2004

LAMBDA: A European System to Access Mathematics with Braille and Audio Synthesis

Waltraud Schweikhardt[1], Cristian Bernareggi[2], Nadine Jessel[3], Benoit Encelle[3], and Margarethe Gut[1]

[1] Universität Stuttgart, Institut für Visualisierung und Interaktive Systeme VIS, Universitätsstrasse 38, 70569 Stuttgart, Germany
{waltraud.schweikhardt, margarethe.gut}@vis.uni-stuttgart.de
http://www.vis.uni-stuttgart.de
[2] Biblioteca di informatica, Università degli Studi di Milano, via Comelico 39/41, I-20135 Milano MI, Italia
cristian.bernareggi@unimi.it
[3] IRIT, 118 route de Narbonne, 31062 Toulouse, France
{baptiste, encelle}@irit.fr

Abstract. This paper deals with an innovative method to improve access to mathematics by blind students in an educational setting. It was designed and developed in the European project LAMBDA (*L*inear *A*ccess to *M*athematics for *B*raille *D*evice and *A*udio-synthesis). This system is made up of two main functional components: the LAMBDA code and the mathematical editor. The research project is funded by the European Union in the framework of the IST Programme. It started in 2002 and will end in 2006.

1 Introduction

It has been a wish since long to facilitate the access to mathematics and science for blind students. Only if blind students can be fully integrated with sighted peers in every learning branch, they can end up with a satisfactory occupation. Three main reasons account for it:

- the lack of a worldwide standard mathematical Braille code, which results in an insufficiency with needed adequate literature,
- the lack of tools, which are specifically designed to aid blind students in manipulating mathematical expressions according to the prerequisites of an educational setting,
- the lack of interactive communication with teachers during the learning process, with colleagues and other scientists.

Traditional mathematical notation for the sighted is international. However, there exist diverse mathematical Braille notations. Therefore, it is hard to share documents among international blind communities.

To give a picture of the situation before the development of LAMBDA we look at the situation in Germany, Italy, and France:

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 1223 – 1230, 2006.

- Germany: In 1930 the Marburg 6-dot notation was specified. It is still in use in German speaking countries, at least in the first four years of education [7]. An 8-dot code for mathematics, SMSB, was available on computers since 1980. Special true type fonts allowed to use it especially in integration [1]. At the university of Karlsruhe the **A**SCII **M**athematik **S**chrift, which uses 8 dots, has been developed [6]. Today, in German schools, also Eurobraille is in use, especially to write down expressions according to LaTeX.
- Italy: mathematics is mainly written on paper using 6-dots Braille. The mathematical Braille code is standardized by Biblioteca Italiana per Ciechi in 1998. An official Italian 8-dots Braille code does not exist.
- France: A commission in charge of the evolution of the French Braille proposed several version for mathematics. The version of 1971 was revised in 2001 and 2006. All versions use 6 dots. Sighted teachers use Bramanet to convert mathematical information written in MathML into Braille [2].

Braille output is achieved in LAMBDA through an 8-dot code, specifically designed to take into account the peculiarities of national Braille codes. Sighted readers can understand what the blind student is writing due to a special font which makes it easy to understand mathematical symbols. Moreover, a traditional two-dimensional representation of the expression is offered in addition.

2 Research and Methodological Approach

User needs' analysis was based both on long years of successful experiences in education for the blind within the LAMBDA working group and on the collection of quantitative data through a preliminary survey. The survey was conducted by means of two questionnaires: the former for the blind students and the latter for the teachers. In each country, the associations for the blind and the universities administered the questionnaire to blind students or teachers, individually. Overall, 121 blind students and 30 teachers answered the questionnaire. This survey fully confirmed the need for a system which enables blind students to write, read, and print text and mathematical expressions as quickly and precisely as sighted peers.

Therefore, the following stages were planned:

- Evaluation of the mathematical editor. About twenty blind students in each country were chosen to test the overall system at school and to perform their daily homework. The evaluation started in 2005 and will end in June 2006, so as to assess the system over one school year.
- Evaluation of the LAMBDA code. The same group of users who tested the mathematical editor was asked to read and understand a list of mathematical expressions chosen from exercise books.
- Evaluation of the font for Braille specific symbols. A group of teachers was chosen to read and understand a list of mathematical expressions linearly presented through this special font.
- Evaluation of conversion tools. Since the LAMBDA system is based on MathML conversion tools, those were tested by using the MathML suite provided by W3C math working group [5].

3 The LAMBDA Code

The LAMBDA system aims to provide the user with full Braille output, therefore, in order to overcome national differences, a set of rules was defined to uniformly describe Braille mathematical expressions in a linear form. Strictly speaking, not a unique mathematical Braille code was defined, rather a common markup language was developed to linearly represent mathematical expressions so that they can be easily read and understood. The LAMBDA project followed the approach, that the blind student can read the mathematical expression in a linear form which preserves as far as possible the peculiarities of the particular national Braille representations. Before the definition of the LAMBDA code, the national codes used in the countries involved in the project were compared and common features were grouped together. The resulting LAMBDA code follows some leading principles:

- Where ever possible the meaning of the linearly represented notation is made explicit.
- It aims to achieve compact linear representations so as to minimize the movements of the finger tips over the Braille display.
- It aims to preserve national dot configurations at least for the most common mathematical symbols.
- Intuitive rules and dot combinations are designed. For example, prefixes are employed to change the meaning of a group of symbols, symmetric dots combinations are used to represent corresponding tags and dots combinations hard to be recognized by touch are avoided.
- Rules which exploit the blank character are not used.
- It aims to be context-free in order to facilitate the design of conversion tools and to reduce the mental workload in understanding the mathematical content [4].

3.1 The Structure of the LAMBDA Code and Its Instancing

Tags may delimit substructures or terminal symbols, namely the leaves of the tree which represent a mathematical expression. Terminal symbols may be single characters, empty tags or also symbols made up of more than one character. A set of acceptable structures (e.g. open-/separator-/close tag) is defined to linearly represent mathematical notations. <open> numerator <separator> denominator <close> describes a fraction. Only three tags are used to linearly mark up this fraction, which achieves a very compact description.

In order to be actually usable in the mathematical editor, the tag structures have to be instanced according to the national peculiarities. This was done for each country involved in the LAMBDA consortium. This instancing process requires:

- the specification of all dot configurations for the symbols which are represented on one Braille cell (e.g. lower-/upper-case Latin letters etc.). These dot combinations have to satisfy the leading principles of the LAMBDA code.
- the definition of the notations to be described. All the mathematical notations used in an educational curriculum have to be linearly described.
- the assignment to each tag of a name to be displayed in the list of tags in the mathematical editor. The name depends on the national languages.

– the assignment to each tag of at least two names to be used in the preparation of the speech output. They are necessary to process the string to be read by the speech synthesizer. At present, two speech reading modes exist. For example, a tag marking the beginning of a fraction can be concisely named "fraction" or, more eloquently, "open fraction".

Each national instance of the LAMBDA code was recorded as XML structure in order to enable the mathematical editor to easily retrieve information and set the local working environment.

3.2 Code Extensibility

One of the most important features of the LAMBDA code is the possibility to linearly describe new mathematical notations. Branches of mathematics incessantly introduce new notations to better express and work with concepts. Some of these notations remain almost unknown and are used by small groups of researchers, but many others get widespread and are commonly taught in the educational curriculum. Therefore, blind students may come across totally new notations without a Braille and speech linear representation. The new notation is linearly described through a tag structure which has to match one of the acceptable structures. Dot combinations and the names to be output by speech are assigned to each tag. To facilitate this process a tool was developed to be used in conjunction with the mathematical editor. It drives the user step by step to modify or extend notation descriptions.

4 The Mathematical Editor

Together with the LAMBDA code, the mathematical editor plays a key role in the LAMBDA system. Its user interface combines all the functional components, namely the LAMBDA code, the modules to handle the layout of the working window and the speech output, the LAMBDA font, the conversion tools, the viewers and the screen readers.

When the editor starts up, a multiple document user interface is presented in the working window. Three main components form the working window: a traditional menu bar, a status bar and an editing area. The status bar displays in particular, messages about the mathematical element pointed by the focus. The editing area is the partition of the window where the user can input text and mathematical symbols. It has to handle multiple not so common operative modes, specifically devised to improve the input, the editing and the exploration of linearly described mathematical expressions. In particular the following functions are implemented:

– A word-wrap mode. It is especially useful to find specific parts in a certain expression quickly, because the navigation can exploit both horizontal and vertical movement commands, which can be input through computer keyboard or Braille display. The most relevant drawback is the lack of comparability between two expressions. Actually, as two expressions are displayed along multiple lines, comparison symbol by symbol is hard to be achieved. When wordwrap mode is not set, the symbols are displayed exclusively along the same line until a new line character is input. In this instance, two expressions can be easily compared symbol by symbol. For example:

```
3*[-(5a-2b)+9*(9a+5a)-2b*(a+1)]=
3*[-(5a-2b)+9*    14a-2b*(a+1)]=
```
(1)

– Skipping of whole mathematical elements which are displayed over more than one Braille cell.
– The possibility to move the focus at any place in the editing area. It means that the focus can be moved directly to a specific position also after the end of line and the end of file. This operative mode is supposed to be particularly effective when a certain two-dimensional layout has to be preserved (e.g. to arrange columnar layouts to solve arithmetic operations). It is often used in conjunction with the automatic indentation mode, which automatically aligns the focus position to the beginning of the previous line when a new line character is input. These operative strategies are useful mainly for Braille display users, because they can better control the spatial relations among the displayed objects.

4.1 Input and Exploration Strategies

Input strategies take into account the necessity to write quickly, to avoid typing mistakes as far as possible, and to reduce the mental workload to remember how to input the desired symbol. Menu based selection, mnemonic short-cut keys, and input based on automatic completion of tag names are available techniques. Furthermore, context based input is designed. The user can choose by a short-cut key whether to input a separation or a closing tag. The LAMBDA editor inserts the proper tag according to the most recently input corresponding tags. For example, this allows to properly close nested parentheses by pressing just one short-cut key multiple times, which was regarded as extremely effective by most users.

Tactile and speech reading usually prevents the blind from getting an overall glance at the mathematical expression and from immediately accessing precise parts inside an expression. Exploration through movement operations means that many operations are provided to move the focus to precise positions in the expression. For example, when the focus is pointing at an open tag, it can be automatically moved to the corresponding separator or to the corresponding closing tag. Exploration looking through the expression structure means, the tag structure of an expression may contribute to understand the overall meaning of the expression and it may help the user to find specific parts. Let us consider the linear form of the mathematical expression:

$$\frac{(x+1)^2}{x} \quad (2)$$

Its linear representation is:

```
//(x+1)^2/x\\
```
(3)

The outermost tag structure is: <open fraction><separator><close fraction>. One step inward, the tag structure is: <open fraction> ()^2 <separator><close fraction>

In tag structure mode, the mathematical editor enables the user to explore the tag structure of a mathematical expression. The building blocks to hide are determined according to the focus position. When the tag structure is displayed, the user can move the focus to a new subexpression so that when the tag structure mode is exited, the new focus position in the actual mathematical expression is the one set in tag structure mode. The main advantage resulting from this operating mode concerns the possibility of getting an overall glance at the linear structure of the expression by reading the building blocks, and, at the same time, to read the symbols actually forming the expression in the unfolded building blocks. Furthermore, it can be used to retrieve and immediately access specific subexpressions without reading every symbol from left to right.

4.2 Editing, Selection Strategies and Speech Feedback

Specific operations were designed to handle mathematical expressions. For example it is possible to delete the content enclosed within a pair of tags, to delete the argument of a function, and so on. Furthermore, selection strategies allow to select the content of tag structures and to extend or reduce the selection to inner or outer building blocks.

The speech output can be provided either through the speech synthesizer driven by the screen reader or through a Microsoft SAPI compliant speech synthesizer. Multiple output modes are available. For example, shortened tag names are used to generate a concise output, generally useful for expert readers, instead full tag names can be used for a more verbose speech rendering. Two studies by Steven [3] and Raman [4] document this.

5 The MathML Export and Import

Concerning the manipulation of mathematical expressions, sighted persons often use graphical applications such as WYSIWYG(*What You See Is What You Get*) software that is focused on the representation of mathematical expressions using the traditional notation. With this kind of software, mathematical information is serialized using a presentation-oriented format (Presentation MathML) to describe precisely the author's presentation intents, leaving to the readers the semantic interpretation.

In order to have a more compact encoding of information in a linear representation, LAMBDA prefers to encode mathematical meaning directly, skipping presentation intents. As a result, mathematical applications dedicated to non-sighted people serialize mathematical expression in a semantic format (Content MathML). The LAMBDA code is directly inspired from Content MathML. As a consequence, conversions between these two kinds of formats (Presentation/Content) are needed for communicating documents between sighted and non sighted persons.

5.1 Import and Export Formats Requirements and Strategies

Our goal was to reuse existing conversion tools in order to minimize coding task in developing only missing conversion processes. As a result, we introduced the Content MathML format as an import/export bridge towards other formats (see Fig. 1).

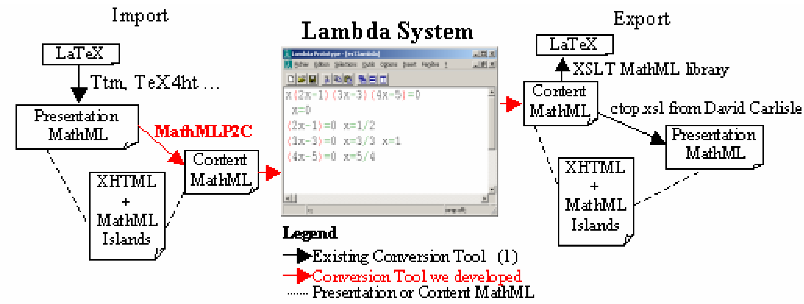


Fig. 1. Import and export format strategy

Many third party conversion tools already exist. We concentrated our effort on the development of the conversion process form Presentation MathML to Content MathML, named MathMLP2C.

5.2 Conversion from Presentation MathML to Content MathML

In general, converting from a presentation oriented format to a content oriented one is difficult to do fully automatically. The main reason is due to the fact that one presentation could be associated to several concepts, depending on the context. For instance, is the expression $a + b$ an arithmetic addition operation or a logical 'or' operation, etc. One possible solution is to specify the mathematical context of information before the conversion.

In order to have a context-aware conversion tool, we gathered mathematical conversion functions into conversion units. One conversion unit corresponds to one mathematical field (context). The following units were developed: Arithmetic, Calculus, Geometry, Linear Algebra, Logic, Set Theory and Statistic.

In order to improve correctness of Presentation to Content MathML document conversion, the user may indicate, before the conversion, the list of units (mathematical contexts) that are involved in the document using a priority order.

6 Conclusions

The LAMBDA system was designed to fit specific user needs arising from groups of blind students and sighted assistants. Therefore, it implements especially new working strategies which aim to compensate for the difficulties due to tactile and auditive perception in getting and managing a proper mental image of mathematical expressions. Furthermore, the need for improved performance in writing mathematics suggested new input, selection and editing. The use of a special font to show the linearly represented mathematical expression and the viewers for traditional mathematical notation close the gap in the communication process between student and sighted teacher.

Thanks to the full integration between the LAMBDA code and the mathematical editor, a uniform working environment is presented to the user so as to reduce as far as possible the necessity to learn different working paradigms and user interfaces, which usually confuses very young students. Also well known successful working techniques and extensibility features are implemented so that the LAMBDA system can be used by deeply different groups of users.

In the end, although the evaluation process is still incomplete, those students who have used it at school for one year, report great satisfaction with the user interface. They remark the need for full integration with mainstream programs (e.g. MathType) which will be completely achieved through MathML within the end of the project.

References

1. Schweikhardt, W. 8-Dot-Braille for Writing Reading and Printing Text which include Mathematical Characters in: Edvards, Alistair D. N., Arato, András, Zagler, Wolfgang L. (ed.) Computers and assistive technology IFIP ´98
2. Miesenberger K., Evaluation of software for the BRAMANET project, Proceedings of the ICCHCP, Linz, 2000
3. Robert David Stevens, Principles for the Design of Auditory Interfaces to Present Complex Information to Blind People, The Department of Computer Science, University of York, 1996
4. Raman, T. V., Audio systems for technical reading, Department of computer science, Cornell University, 1994
5. W3C voice browser working group, Speech Synthesis Markup Language (SSML): Version 1.0, World Wide Web Consortium, = {2004 W3C Math Working Group, Mathematical Markup Language (MathML): Version 2.0, World Wide Web Consortium, 2003
6. Klaus, J.: Studium für Sehgeschädigte in Informatik und Wirtschaftsingeni-eurwesen. Abschlußbericht des Modellversuchs Informatik für Blinde. Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie. Bonn 1995
7. Hertlein, J. (Hrsg.): Internationale Mathematikschrift für Blinde. Marburger Systematiken der Blindenschrift. Teil 6. Deutsche Blindenstudienanstalt. Marburg 1992

Speech Recognition Helps Visually Impaired People Writing Mathematical Formulas*

Tomáš Hanakovič and Marek Nagy

Department of Applied Informatics, Faculty of Mathematics, Physics and Informatics, Comenius University, Slovak Republic
hanax.s@gmail.com, mnagy@ii.fmph.uniba.sk

Abstract. We present a web-based application that can use a speech dialog to write mathematical formulas. It is suitable also for visually impaired people. A voice control is simple and easy-to-use to reach accessibility goals. All functions are separated to relevant categories in menu and accessible via mouse as well as voice. Application is based on X+V [1] technology, java web-server and browser-side scripting.

1 Introduction

Mathematical expressions are basic components in the domains of math, physics, computer science and engineering. People can write them on a computer in many ways utilizing various applications. For example very popular mathematical editor TeX [2] or WYSIWYG editors are used. In this contribution we present a web-based application, which can be used as an accessible formula editor for visually impaired people as well as a simple application interface for the rest. The main principle is to use a speech dialog to enter mathematical formulas. A voice control is simple and easy-to-use to reach accessibility goals.

1.2 Goal

The goal is to make a formula writer, which is suitable for healthy as well as disabled people. An application should be as much intuitive as possible in order to minimize confusion and maximize efficiency (especially working speed). It will be good to apply existing software technologies to increase compatibility and reliability.

2 Analysis

The most important factor is intuitiveness, but this requirement has vague definition and needs to be analyzed. Application with intuitive control should meet several categories of needs. We can formulate at least two of them: usability and accessibility.

2.1 Usability Needs

There are several points that should be reached to improve application usability. At first we need a simple graphical user interface (GUI), so that user is not confused with

* Partially supported by a national grant VEGA 1/1055/04.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 1231 – 1234, 2006.

many controls and need not think too long about “How to do that?”. However, it is necessary to offer nontrivial set of functionality. These two requirements are opposite therefore solution must be somewhere in the middle.

Next, we need a good orientation in the formula. It consists of ability to quickly identify chosen element and navigate to it. In practical use we do not need to choose and identify every atomic element in formula. We only need to locate a structural pattern where an element is located and then identify its sub-elements. Thus navigation should be done hierarchically.

2.2 Accessibility Needs

Instead of using GUI, visually impaired user is required to use other modalities. We have chosen voice because of aforementioned intuitiveness and straightforward usage. However, existing voice technologies are not able to simulate human semantic recognition abilities, so several restrictions must be applied to user input. In mathematics there are several well-known and commonly used spoken constructs, so it will be good to match them as close as possible.

The main problem, which must be solved, is a difference between exactness of mathematical expressions and intuitive reading of them. Many common verbal constructs are not formed exactly and they lack exact interpretation. Readers often ignore syntactic elements which are implied by the context (e. g. *“a i k”* can be interpreted as ”a_{ik}“ as well as “$a_i k$“). It can be solved by including a strict syntax. It is not comfortable but it provides straightforward way to voice dictation. Of course, the syntax should be as short as possible but still expressive.

Regarding to orientation in the formula, this can be serious problem for visually impaired user, especially if formula is too complex. It's desirable to have some intuitive element numbering as well as voice read-out with suitable complexity and verbosity. While navigating, visually impaired user should be aware of context of selected element, just like healthy people can see enclosing elements. Again, hierarchical approach will be necessary.

3 Main Solution

To reach good usability and accessibility we should arrange formula elements into categories. This approach allows to sort program functionality into visual categories (for example menus). In each category we force to use specific syntax, which minimizes voice recognition error (we have relatively small set of valid inputs in context of specific category) and also supports exact semantic interpretation.
Let us consider typical example in which user wants to compose equation:

$$sin\ 2k + l = {}^{k+l}/_{k}\ . \tag{1}$$

Using commonly used conventions we can for example formulate category *function*, which can contain well-known functions (sinus, logarithm,...). Using similar categories for all formula elements we can split equation into sequence of categories:

function, *linear expression element*, *equation symbol*, *fraction*

Then input sequence (elements separated by semicolon) will look like:

function *sinus*, parameter *2k*; *plus l*; symbol *equals*; fraction, nominator $k + l$, denominator k

It can be noticed that we can do this categorization in several levels of formula (for example we can have function in the fraction nominator). This observation indicates that we use aforementioned hierarchical approach. Above technique can also be applied to numbering elements: we index first level elements sequentially, choose structure we want to expand and then index selected structure at second level. We repeat from second step until chosen element is reached.

Secondly, we use prosody to minimize reading of structural aspects in voice read-out. Above equation will be read as *"sinus of two k <pause> plus l equals fraction k + l over l"*. Note how *<pause>* element simply indicates end of sinus parameter.

4 Design

Choosing suitable software can simplify and speed up development process and it also determines set of possible functionality. With regards to recent trends and technologies we decided to use Java language and X+V [1]. This choice allows to control voice processing and application logic separately. Final output from application is realized in the presentation MathML [3], which was chosen with respect to existing standards.

4.1 Application Logic

Besides initialization phase, main application insert mode consists of several steps sequentially processed in execution loop. At first user is prompted (visually and/or by voice) for entering (mouse or voice) input, which is locally validated and sent for next processing. Because of linear nature of voice input (sequence of elements) we need to delinearize it and restore original structure[1], which involves identifying above mentioned categories (variables, operators, functions, symbols, etc...). This structure is then inserted into expression tree that represents entered formula. This tree is main data structure composed in successive steps and all subsequent operations are done on it, namely generating visual, voice and MathML output to the user interface.

Application also offers working in other modes, namely navigation and edit mode. Navigation mode allows to identify or search for certain structure and edit mode makes corrections possible.

4.2 Technologies

From a technical point of view, the application can be represented as a two-layer system. Bottom layer (application core) is made up from java [4] web-application that runs on Jetty [5]. Top layer (audiovisual interface) consists of speech recognition and rendering mechanism and GUI. Opera browser® [6] was chosen because it

[1] Input sequence *"a plus b over c"* should be restored into form *"*$a + {}^{b}/_{c}$*"*.

successfully supports X+V technology, has native speech recognition and rendering engine and as web browser it can serve as suitable GUI.

A java web-application offers two main services. First one is parsing input and building expression tree and the second one is generating X+V pages for browser.

Audiovisual interface is represented by X+V pages which contain visual (XHTML [7]) and voice (VoiceXML [8]) part. Besides displaying output, X+V page also offers functionality to gather an audio input from user. This can be done by short voice conversations or manually by selecting certain menu or visual element.

We use SSML [9] to format prosodic characteristics of speech output.

References

1. XHTML+Voice Profile 1.0, [http://www.w3.org/TR/xhtml+voice/]
2. TeX Users Group, [http://www.tug.org/]
3. Mathematical Markup Language (MathML) Version 2.0 (Second Edition), [http://www.w3.org/TR/MathML2/]
4. Java Technology, [http://java.sun.com/]
5. Jetty, [http://jetty.mortbay.org/jetty/index.html]
6. Opera browser, [http://www.opera.com/]
7. XHTML (tm) 1.0 The Extensible HyperText Markup Language (Second Edition), [http://www.w3.org/TR/xhtml1/]
8. Voice Extensible Markup Language (VoiceXML) Version 2.0, [http://www.w3.org/TR/voicexml20/]
9. Speech Synthesis Markup Language (SSML) Version 1.0, http://www.w3.org/TR/speech-synthesis/]

Supporting Blind Students in Navigation and Manipulation of Mathematical Expressions: Basic Requirements and Strategies

Bernhard Stoeger, Mario Batusic, Klaus Miesenberger, and Philipp Haindl

Institute of Integrated Study, Johannes Kepler University Linz
Altenbergerstrasse 69, 4020 Linz, Austria
{bernhard.stoeger, mario.batusic, k.miesenberger, philipp.haindl}@jku.at

Abstract. In [10], the problems faced by a blind or visually impaired student in doing Mathematics were analyzed, and the basic ideas of a MAWEN (Mathematical Working Environment), a software solution to help overcome these pressing difficulties, were described. The present paper builds upon the latter one, refining the ideas sketched there. After a thorough description of the state of the art, we present some general considerations on the problems met by a blind pupil when navigating within mathematical expressions and when doing calculations. Finally, through several case studies taken from mainstream school books, strategies to provide computer aided support to overcome the problems are outlined.

1 Introduction

In [10], four basic problems faced by a blind or visually impaired student working in Mathematics were formulated:

- Access to mathematical literature (books, teaching materials, papers etc.)
- Preparation of mathematical information (presenting school exercises, writing papers etc.)
- Navigation in mathematical expressions
- Actually doing Mathematics (carrying out calculations and computations at all levels, doing formal manipulation, solving exercises)

A fifth basic problem, which is closely related to all the above mentioned, is to be added: The problem of mathematical communication between sighted and blind/visually impaired people, typically, communication between a sighted teacher and a blind/visually impaired pupil.

Whereas some work towards the first and second problems done in [8] led to practical solutions, problems 3 to 5, although touched by several research initiatives, are still waiting for solutions applicable in practice.

Our idea to design a comprehensive multi-modal software solution to address the above listed problems, especially problems 3 to 5, which we call a MAWEN (Mathematical Working Environment), was outlined in [10]. the present paper is to refine the ideas set forth there, by:

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 1235–1242, 2006.

- analyzing current mathematical school book literature, in order to isolate some of the implicit methodology used by mainstream mathematical teaching;
- providing some more case studies than in [10]
- discussing several key problems in navigation through complex mathematical expressions
- describing several standard mathematical exercises, outlining the problems met by a blind or visually impaired pupil in doing them, and discussing supportive tools to perform these exercises, to be integrated in a MAWEN

2 State of the Art

One of the first steps to be taken to mathematically enable blind people was the creation of so-called "mathematical notations". Most of them are codes that build upon traditional Braille [12] to [9]. They have the advantage that they are quite well optimized for a Braille user, but they do not permit easy sharing of information between blind and sighted people. One of their further disadvantages is the fact that a very large number of such codes have been developed over time - every attempt towards a unification or standardization of Braille maths codes failed by now.

In order to make communication somewhat easier, several notations based on ASCII were defined [13]. Such ASCII maths notations enable communication between blind and sighted people through a compromise: Whereas the sighted partner does not see a formula in its familiar graphical rendering, the blind partner gives up the optimization of maths notation towards Braille. On the other hand, most ASCII maths notations are easy to learn, so they are practicable to some extent.

One special case of an ASCII notation is the source code of the TeX typesetting system, especially that of the widely used LaTeX macro package. Since, at least after some revision, mathematical documents in LaTeX source are accessible to the blind, LaTeX offers a partial solution to basic problem number 1. This is, to a limited extent, also true for problem number 2, since a student who writes his/her exercises in LaTeX source needs just to invoke the LaTeX compiler in order to have a well-formatted and very attractive hard copy of his/her work for the sighted teacher.

In order to get more practical solutions out of braille maths codes, ASCII math codes, and LaTeX source, two initiatives were started: the LabraDoor software package [8] developed by our research group converts LaTeX documents into the Marbourg System [3], one of the most widely used Braille codes in German speaking countries, or into HRTeX, a new notation that resembles TeX, but eliminates some of that code's more awkward aspects. LabraDoor is now the basis for the production of school books for blind pupils in Austria, largely overcoming problems 1 and 2 on a primary, secondary, and high school level.

The second initiative comes from the University of Pierre and Marie Curie in Paris. It is called UMCL [1], which means "Universal Math Conversion Library".

It is a C API for converting various maths notations between each other, together with several concrete conversions already implemented. In the end, UMCL shall allow every partner involved in mathematical communication, blind or sighted, to use the notation of his/her choice through the UMCL and the applications designed for it shall do the necessary conversions in the background, via an internal representation of maths expressions following the MathML standard from W3C [14].

Whereas problems 1 and 2 will be largely overcome by the initiatives listed above, much more work has to be done towards problems 3 to 5. One first step addressing problem 3 was taken by the University of South Florida, who are developing the MathGenie, a tool that enables one to browse through a mathematical expression by synthetic speech [6]. It is one of the goals of MAWEN to implement navigation support that is to work for a Braille user as well.

A tool that might support a pupil both in navigation and manipulation, i.e., would address problems 3 and 4, is the Virtual Pencil, developed by the company Henter Math, LLC [5]. It addresses the needs of people who are unable to operate a pencil effectively, which is a key part in learning and dealing with maths. As it virtually takes on the role of a pencil while at the same time avoiding an overload of the student with unsolicited educational support it is an appropriate solution for in-class maths education.

A further project on access to Mathematics is Lambda [7], which provides a braille and audio device based system for the management of scientific documents. Hereby the focus clearly is put on a widely customizable user interface, which allows the pupils to edit mathematical notations as well as supports the learning progress by offering different interfaces for team and software collaboration for interactive maths education.

A further project on the topic is Infty [4], which is dedicated to the development of a software for mathematical OCR. They also build a tool to input and to edit mathematical formulae based on voice output.

Finally, we would like to mention the EU-funded project MICOLE [2], which means "Multimodal Environment for Inclusion of Visually Impaired Children". Much work undertaken by us for the MAWEN is being carried out within MICOLE, the MAWEN being one component of the environment to be designed within this project.

3 Research Methodology

In order to supply usable support of a blind individual in his/her mathematical tasks, it is of key importance to gain an understanding of the methods currently exercised by blind people working in Mathematics. As begun in [10], these methods have to be analyzed, exposing the problems, and inefficiencies, presently met when applying them. Hereafter, suggestions to improve the methods in order to better meet the needs of the target group have to be worked out.

Since present mathematical teaching, also that applied to blind pupils, is strongly influenced by visual presentation, the programme outlined above will need an understanding of the methodology currently applied by sighted people when doing mathematical exercises, methodology currently conveyed by today's mathematical teaching. It is strange that such methodology appears not to be reasonably documented in the literature. Rather, the methodology seems to be implicit in teaching, especially in school books. We are therefore analyzing some of these books, in order to isolate that implicit methodology. This analysis shall lead to a deeper understanding of the way how Mathematics is presently done. It shall be the basis for the formulation of a set of functions, also called support or helper routines, to be implemented into a software in order to aid a blind or visually impaired student in his/her mathematical exercises. In the MAWEN, these functions shall be implemented, and some of them shall be combined into wizards, as shown, e.g., in our MICOLE prototypes [2].

4 Findings

4.1 Considerations on Text Selection

When dealing with mathematical formulae on the computer, it is very often necessary to select, or mark, a piece of an expression for further use. As an example, you might have a complex expression in parentheses, which you need to simplify prior to being able to continue your calculation. you will then want to start a Simplification Assistant, as outlined in [10], to be applied to just that particular expression in parentheses. To tell the assistant that it should handle the expression, you need to mark it in your working environment.

Of course, you could do this by standard GUI methods - traditional keyboard shortcuts, usually involving the Shift key, or mouse actions, typically Drag and Drop. But this procedure is not always what a blind pupil would desire: For first it is not always easy to navigate within a selected portion of text; in particular, you cannot use the cursor in such a case, since selection will be destroyed as soon as you move the cursor. Secondly, you want to have a possibility to mark particular expressions in a quick way, not involving a walk through the whole expression. To remain with the above example of a parenthesized expression to be selected, it would be desirable to have a function that, when pointing at an opening parenthesis, automatically selects everything between that parenthesis and the matching closing parenthesis.

The first of these two arguments, the problem that the cursor cannot be used when something is selected, would inspire the idea to create a concept of selection that replaces the one inherited from standard GUI, a type of selection that is independent from the cursor. Within such a setting, one could facilitate selection of a portion of text by just pointing at the start and at the end position of the text to be selected. The second argument, the desire for automatic selection tools based on the structure of an expression, would motivate routines such as:

- Selecting the current summand of a sum,
- Selecting the current factor of a product,
- Selecting the current side of an equation, etc.

4.2 Considerations for Navigation Support (Problem 3)

One of the simplest aspects of navigation support is the ability to collapse and expand an expression, similar to what is offered by modern programming environments: When confronted with a very long and complex formula, the user should be able to temporarily hide portions of it from the display. At a later time, he or she should be able to make these concealed pieces visible again. Such a technique poses several questions, e.g.:

– How should the user specify the portion to be concealed?
– By which means - braille characters etc. - should the user be informed about the presence of a concealed piece in a formula?
– By which means should expansion of a collapsed portion be commanded

When examining these questions, it becomes apparent that one needs to distinguish between representational and structural view of a formula: If a long and complex expression is presented, it might be desirable to specify a piece to be collapsed by just selecting its start and end position within the formula. However, such a technique, although fairly easy to implement, would not prove too efficient when the piece to collapse gets lengthy: What one might desire would be a condensed view of an expression, telling the user only coarse structural information about it. To take an example: When dealing with an intricate fraction, a condensed view might only convey the information that a numerator, a fraction line, and a denominator are there. Should the user need more information about, say, the numerator, a command needs to be implemented to expand that part of the expression, but, perhaps, not to full depth.

Another family of helpful functions might be what we call path support: Inside an expression, clicking at a spot should give detailed information on the path leading to that spot within the structural tree of the formula. This could be conveyed by a path string similar to a path name in a file system. As an additional support feature, clicking at a branch in such a path string could display the subexpression corresponding to that branch.

4.3 Considerations for Manipulation Support (Problem 4)

In order to aid blind students in formal manipulation, i.e., in doing standard mathematical exercises, elaborate, efficient, and flexible navigation support is essential. In addition, the main classes of mathematical tasks met in education need to be identified - this will be done through our analysis of school books announced in the previous section. Such standard tasks might be:

– Multiplication of sums in parentheses [10]
– Simplification of a sum [10]
– Solving a linear equation
– Working with fractions
– Doing basic calculations in arithmetics
– Multiplying and dividing polynomials
– Doing elementary vector and matrix operations
– Solving systems of linear equations

Through analyzing tasks of that kind, a set of standard support functions shall be isolated. The ideal situation would be to implement this set of standard functions, such that support for every standard task like the above may easily be given by combining several of those functions, either in terms of a wizard, or in terms of an environment that lets a student choose freely which support functions he or she would like to use to perform a particular task.

In what follows, we shall present several aspects of manipulation based on standard exercises taken from a school book. For every such aspect, the problems it poses for a blind pupil are considered, and ideas to give support in overcoming the problems are outlined.

Support for Side Calculations, and Calculator. In an exercise like 559,c in [11]:

$$7,1x + 4,8x - 0,5x =$$

you need to add the numbers $7,1$, $4,8$, and $-0,5$ in order to get the coefficient of the variable x. You should not be expected to do such an addition in your memory. Rather, you should be able to organize it in terms of a side calculation: The working environment should allow you to open a second instance of itself, with the sum $7,1+4,8-0,5$ as input. This input could be constructed by walking through the original exercise, marking each of the three numeric summands, and issuing a command that presents the three marked numbers as input values of a side calculation, whose type is Addition. From that point, the side calculation could be solved either by starting the arithmetic addition assistant described in [10], or, in case of a more advanced teaching setting where the use of a calculator is permitted, by invoking a calculator emulation built into the MAWEN.

In any case, side calculations should be easily identified for further reference, which could be done by assigning names to them, or by automatically numbering them.

Support for Parentheses; Passive Assistants. In an exercise like 587,a from [11]:

$$5a - 3b - (2a - 4b + 8) + (2b + 4)$$

the second parenthesized term is no problem, because the parentheses can simply be omitted. However, the first parenthesized term needs consideration, because resolving the parentheses involves swapping of sign in every summand. As always in such situations, it would be inconvenient for the pupil to be required to remember every summand in order to copy it into the interim result with swapped sign. A convenient way of support would be furnished as follows: first, an identical copy of the input string is made, connected to the input through an = sign:

$$5a - 3b - (2a - 4b + 8) + (2b + 4) = 5a - 3b - (2a - 4b + 8) + (2b + 4)$$

Then, the pupil walks through the right side of the equation, deleting the opening parenthesis after the $-$ sign, manually inverting the sign of the two remaining summands, $-4b$ and $+8$, and deleting the closing parenthesis. In such a way, the parentheses are resolved conveniently. We would like to call such a kind of assistant a passive assistant, because no active software support is given - the pupil does everything that has to be done manually, but still without inconvenient load on his/her memory. The passive assistant could be made partially active if it marks every sub-term whose sign has been manually swapped by a checked sign, but such high-level support should be optional.

In many cases, parentheses are nested, just like in Exercise 589,a from [11]:

$$6s - [3s + 2t - (s + 2t)]$$

Here, one must first resolve the inner parentheses, such that the contents of the outer ones, which are here written as square brackets, can be simplified. After doing so, the whole expression must be simplified again.

This is a typical example for a situation where a sub-calculation should be initiated: First select the contents of the square brackets, in order to start a Simplification Assistant with the selected expression as input. Within this sub-calculation, a passive assistant to resolve the inner parentheses, as described above, may be used. Once the expression, which originally was the contents of the outer parentheses, has been successfully simplified, it is time to return from the sub-calculation into the main one, with the contents of the outer parentheses replaced by their simplified version. What remains to be done is a simplification of an expression without parentheses.

Doing Tests for an Algebraic Calculation. In elementary algebraic teaching, you are often required to test the correctness of a manipulation by assigning concrete numeric values to a variable. As an example, Exercise 566,a in [11] instructs you to simplify the expression,

$$6a + 4b - a + 5b - 3a$$

which results in $2a + 9b$, and to test the correctness of your calculation by assigning the value 2 to the variable ”a$and3to$”b.

For a blind pupil to carry out such a test, one problem will arise, namely, to remember the values, in our case 2 and 3, to be assigned to the variables. This could be supported by a kind of passive assistant: the pupil copies both expressions whose equality is to be tested, the original input and the result, into an edit buffer. Now, in order to compute the numerical value of one of these expressions resulting when the variables are assigned to numbers, the pupil will walk through the expression, replacing every occurrence of the variables by the numbers. this can be done by simple editing - it is a passive assistant, leaving the work to the pupil. However, in order to give support in remembering the number to be used instead of a variable, the assistant could prompt that number whenever the pupil points to the corresponding variable.

References

1. D. Archambault. UMCL Homepage. On-line resource: http:\\inova.snv.jussieu.fr\iGuma\umcl\, Apr 2006. (accessed 11 Apr 2006).
2. University of Tampere Finland Department of Computer Sciences, Tampere Unit for Computer-Human Interaction (TAUCHI). The MICOLE Project. On-line resource: http:\\micole.cs.uta.fi, Apr 2006. (accessed 11 Apr 2006).
3. H. Epheser, D. Prograniczna, K. Britz, and B. Betz-Weitzel. *Neufassung und Vervollstaendigung des Systems der internationalen Mathematikschrift fuer Blinde. Marburger Systematiken der Blindenschrift.* Number 6. Deutsche Blindenstudienanstalt, 1986.
4. Faculty of Mathematics and Graduate School of Mathematics, Kyushu University, Japan. Infty Project Homepage. On-line resource: http:\\www.inftyproject.org\, Apr 2006. (accessed 11 Apr 2006).
5. Henter Math LLC. Virtual Pencil Homepage. On-line resource: http:\\www.hentermath.com\, Apr 2006. (accessed 11 Apr 2006).
6. A. Karshmer. MathGenie Homepage. On-line resource: http:\\karshmer.lklnd.usf.edu\~mathgenie\, Apr 2006. (accessed 11 Apr 2006).
7. Lambda Consortium. Lamda Project Homepage. On-line resource: http:\\www.lambdaproject.org\, Apr 2006. (accessed 11 Apr 2006).
8. K. Miesenberger, M. Batusic, and M. Stoeger. LABRADOOR, a contribution to making mathematics accessible for the blind. In A. Edwards, N. Arato, and W. Zagler, editors, *Computers and Assistive Technology - 6th International Conference on Computers Helping People with Special Needs ICCHP 98.* Oldenbourg, Wien, Muenchen, 1998.
9. A. Nemeth. *The Nemeth Braille Code for Mathematics and Science Notation: 1972 Revision.* American Printing House For the Blind, New York, 1972.
10. B. Stoeger, K. Miesenberger, and M. Batusic. *Multimodal Interaction with Mathematical Formulae: Access, Communication and Support for Blind People in Doing Mathematics.* L. Lawrence Erlbaum Associates, Inc. (LEA), 2005.
11. W. M. Stutzka. *Lebendige Mathematik 3*, volume 1. OEBV/HPT Wien, 2001. ISBN 3-209-03756-6.
12. The Shodor Education Foundation Inc. BRL: Braille Through Remote Learning. On-line resource: http:\\www.brl.org\, Apr 2006. (accessed 11 Apr 2006).
13. Universitaet Karlsruhe, Studienzentrum fuer Sehgeschaedigte, Karlsruhe. ASCII-Mathematikschrift Karlsruhe/Dresden. On-line resource: http:\\web.inf.tu-dresden.de\~js28\material\ams\h-000001.htm, Apr 2006. (accessed 11 Apr 2006).
14. World Wide Web Consortium. MathML Homepage. On-line resource: http:\\www.w3.org\Math\, Apr 2006. (accessed 11 Apr 2006).

Scientific Diagrams Made Easy with IVEO™

John A. Gardner and Vladimir Bulatov

ViewPlus Technologies, Inc, 1853 SW Airport Ave, Corvallis, OR 97330 USA
{john.gardner, vladimir.bulatov}@viewplus.com
http://www.viewplus.com

Abstract. Virtually all modern scientific documents and textbooks use graphical illustrations and/or data displays. The ViewPlus IVEO™ technology, based on Scalable Vector Graphics (SVG) is being developed to permit scientists to publish such graphics in a form fully usable by all people. The first release of IVEO™ was designed to make simple graphics accessible. Development of the second release, which includes improved text accessibility, linking, and interactivity, is described in this paper. We also discuss new research to expand accessibility of quantitative data in IVEO™ documents by using non-speech audio.

1 Introduction

Charts and diagrams are essential in both pedagogical and professional scientific literature. They are typically "made accessible" to blind people either by word descriptions or by tactile diagrams. Word descriptions are also conventionally used to assist people with other print disabilities to comprehend graphical information. Neither word description nor tactile graphic representation is fully satisfactory. Scientists would not use diagrams if word descriptions were equally informative. An English proverb is that "a picture is worth a thousand words". Indeed, many scientific pictures could not be described adequately with many thousands of words.

A tactile diagram can be very useful to a blind person. Unfortunately very few blind people are adept at reading tactile diagrams. Opinions differ on why tactile diagrams are difficult to read, but most people would agree that one needs to learn to read tactile diagrams, and few blind people have had any opportunity to learn and practice that skill. The tactile sense has much less spatial resolution than normal sight, so transcribers must enlarge and simplify most published graphically displayed information in order to make it comprehensible to a blind reader [1].

The concept of using both tactile and linked audio information was pioneered approximately two decades ago by Parkes [2-6] with the Nomad concept. Blind users felt tactile graphics that were mounted on a touch-sensitive pad connected to a computer. When one pressed in various regions, the computer provided speech feedback to identify the object and then permitted the user to bore in for additional information. Nomad was used in a number of commercial products [7,8] and hailed as a major advance in "making graphics accessible", but creation of tactile copy and the associated computer audio information was quite laborious. Consequently Nomad has been used largely for special sets of curricular diagrams prepared for large numbers of student users, and not to convert scientific graphics to accessible form for single users "on request". Most mainstream scientific literature has remained solidly inaccessible.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 1243 – 1250, 2006.

The authors of this paper, in collaboration with colleagues at Oregon State University and ViewPlus®, set out systematically to develop the missing technologies that would permit “Nomad accessibility” to mainstream scientific graphics [9-13]. It was clear that one such missing technology had to be a simple method of producing tactile graphics from mainstream pictures that could have sufficient tactile resolution for one to discern features of interest. That research direction led to the Tiger® embossing technology and the popular ViewPlus® Tactile Graphics and Braille embosser products [14,15]. Any graphic that is adequately represented by a gray scale image will be faithfully converted to a tactile copy by embossers employing Tiger® technology. Dark regions are represented by big dots, and light areas by small dots. Generally it is possible to distinguish tactually the features of such graphic pictures even if one may have no idea what all the bumps and waves actually represent.

The second major missing technology was some simple way to create the audio information needed by a blind user to understand the meaning of those bumps and waves. This research direction led to the IVEO™ software applications introduced commercially by ViewPlus® in 2005. The IVEO™ Creator application and an associated utility, the IVEO™ Converter, permit one to create simple Scalable Vector Graphic (SVG) files directly or to scan in images from paper or convert files from other electronic formats to SVG. Authors can easily provide labels and, if desired, long descriptions of graphical objects. SVG permits both title and description fields for graphical objects within the SVG format. When a user of the IVEO™ Viewer application clicks on that object, the label is spoken. Descriptions can be heard by pressing a hot key. Braille users can “hear” these words with an on-line Braille display. The title and description attributes would be useful to a wide variety of users, so it is possible that widespread use of the IVEO™ technology could persuade mainstream authors of the value of taking the small amount of extra time required to insert title and description attributes for important objects. If this becomes routine practice, IVEO™ would achieve the intended goal of providing universally usable graphical information. Many scientific graph, bar/pie chart, and diagram authoring applications could automatically insert object titles as well as quantitative data, making the SVG graphic information-rich as well as accessible.

IVEO™ Creator automatically structures SVG text into geometrically-associated portions that are normally semantically related. When a user presses any of that text, the phrase is spoken. A substantial fraction of scientific diagrams can be understood by a blind person who feels the diagram and reads the labels audibly, so these are automatically made accessible simply by saving from Creator. Good text structure is important, so improvements in the capability of authoring systems to export structured text would be an improvement. IVEO™ Creator currently guesses at the text structure and doesn’t always get it right.

SVG is a new but increasingly popular graphics language [16]. IVEO™ files can be viewed by sighted users with the popular Adobe [A] SVG Viewer or with the IVEO™ Viewer [A]. Both viewers are free and of comparable size. Only the IVEO™ Viewer provides access to blind users. It also has good accessibility features for people with other visual and learning disabilities.

2 IVEO™, Release 1

The first release of the IVEO™ software was intended primarily for access to relatively simple graphics. It was intended to be useful by a small number of blind power users and by agencies that make simple graphics accessible to people with print disabilities. Maps and diagrams with sparse text labels that have been created with standard Windows® graphics author/editing applications can easily be converted to SVG by printing from the original application to IVEO™ Converter. The resulting SVG file is opened in IVEO™ Creator where important graphical objects may be labeled by a transcriber. Long text labels can be checked to assure that the groupings made by Creator are semantically correct. The IVEO™ SVG file is then saved and sent to the end-user by e-mail, CD, or other convenient media. Tactile copy can be made by the transcriber or by the end user. It is made by printing the IVEO™ page to a ViewPlus® embosser. If the end user creates his own copy, he has control of size and various user options for printing. More importantly, the user can zoom features and make additional printouts to access smaller features that may be difficult to understand on the original diagram. Figure 1 is an example of a moderately complex scientific diagram, the solar system that is accessible with IVEO™ release 1. The dots that are embossed when tactile copy is made are illustrated in Figure 2.

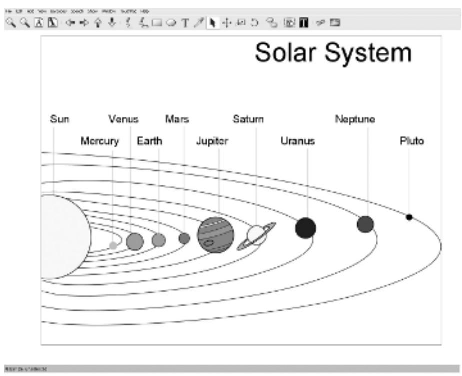


Fig. 1. Screen image of a solar system diagram as displayed with the IVEO™ Viewer application. Several display modes are selectable including an uncluttered full screen "accessibility mode".

When Creator is used by a blind person, the user makes a tactile copy and has access to text labels. Typical diagrams used in science, math, and business are often reasonably accessible without any intervention by transcribers. Examples include line diagrams with explanatory text labels, simple flow diagrams, and many kinds of charts. The diagram shown in Figure 1 can be made accessible simply by converting the original document to an IVEO™ SVG file, creating an embossed copy, and reading the labels while feeling the orbits. A good caption or figure description and additional labels identifying the orbits would make the diagram more rapidly comprehensible, but an experienced blind user should understand diagrams of this complexity with no intervention by a transcriber.

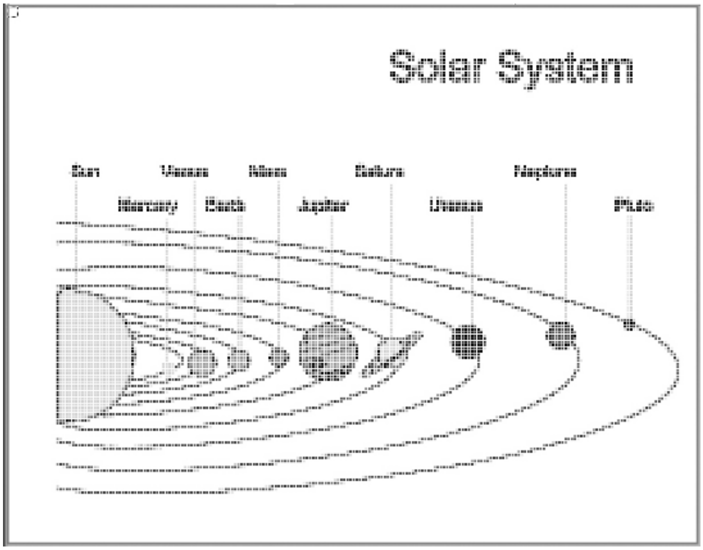


Fig. 2. Illustration of the dots made by printing the file of Fig.1 to a ViewPlus embosser. Dark dots are high, and light dots are low when embossed.

The IVEO™ Viewer also has features intended to enhance access by people with other print disabilities. People with low vision are finding that the customizable status bar that repeats audio information is helpful. It is expected that the same will be true for most people with learning and cognitive disabilities. Work is currently in progress to test such users and to improve IVEO™ display features to make IVEO™ as usable as feasible for them. The new ViewPlus® Emprint™ haptic color printer produces both an embossed and superposed color image that has been long anticipated by professionals working with students having learning disabilities. Emprint™ images used with the IVEO™ Viewer and a touch pad should provide very good access for many people with severe print disabilities, e.g. dyslexia.

The first IVEO™ release permits access to text only by touching a portion of a phrase and hearing it. This is good access for diagrams with small labels, such as Figures 1 and 2, but it is ponderous for charts having large amounts of text in tables, lists, etc. New features for better text access are needed for blind users. Transcribers also need some additional features, and current development for second and future releases are described below.

2.1 Testing

IVEO™ was developed over a number of years during which time, many testers and consultants influenced its development and feature set. Extensive beta testing was done by a small but dedicated group of blind users prior to commercial release. These people suggested some expansion in the feature set but mostly assisted ViewPlus® to find bugs and make the software more robust. This is the appropriate way to test commercial software.

The most important question for IVEO™ developers was whether IVEO™ could provide good Nomad-like access to mainstream information without extensive editing by transcribers. Alpha and beta testing was sufficient to establish that it could. ViewPlus® did not attempt to answer more detailed questions such as percentage of people with print disabilities who could usefully access mainstream literature without assistance, or whether mainstream use of IVEO™ could provide access to all users. These are interesting and worthy questions but are more appropriate to research by disinterested academic researchers than by a company marketing the product.

3 IVEO™ Expansions

3.1 Linking

The first IVEO™ release permits users to hear labels, descriptions, and text in synthesized speech. Synthetic speech is generally considered inferior to recorded human voice, and many students with learning disabilities understand human speech much more easily. Beta testers and early adopters requested that IVEO™ be expanded to permit recordings instead of or in addition to ability to use of synthetic speech. This capability will be included in the upcoming second release.

The capability to play digital recordings is provided by links within the IVEO™ SVG file that open audio applications in the background to play audio files in various file formats. The linking feature has been implemented more broadly so that users can now open up other applications by clicking a link in the file. A common example would be a link to some web page that transfers control to the users default web browser. The user returns to the SVG document when the web page or pages have been read. This link-to-external-files feature is included in the second release. Beta testing had just begun at the time this paper was written, and no results are yet available.

3.2 Interactivity

SVG is a powerful language that permits authors to introduce a rich variety of interactive content. Interactivity was not permitted in the first IVEO™ release but will be introduced in a later release. The next release of the Viewer will support some limited interactive content, but it is largely a research tool for ViewPlus® and academic researchers. Providing accessible interactive content is still very much a subject of research. Authoring tools permitting some limited interactivity may be available soon, but it is expected that adequate accessible authoring applications will take years to develop.

3.3 Improved Text Access

Short text phrases are commonly used on diagrams and maps to identify objects, on graphs to label axes and give titles, and for many other purposes. Blind users can read them easily by pressing on such text. Unless the graphic is considerably enlarged, common text is reproduced by the embosser as a ragged line and is usually easy to distinguish from graphics. However when text is dense it becomes cumbersome to read intuitively by touching it. It is difficult to locate the position of a text phrase in standard sized text. Fingers do not have the resolution of vision, and it is difficult to move right and left or up and down without losing ones place.

Figure 3 illustrates the difficulty of accessing such text with the current IVEO™ touch-and-speak method. The fingers examining the tactile copy have fingertips that are larger than the spacing between the text lines. The second IVEO™ release will permit users to select regions and read the text with the keyboard much like reading text in a word processor. This development is still in early stage when this paper is written, and beta test information is not expected for some time.

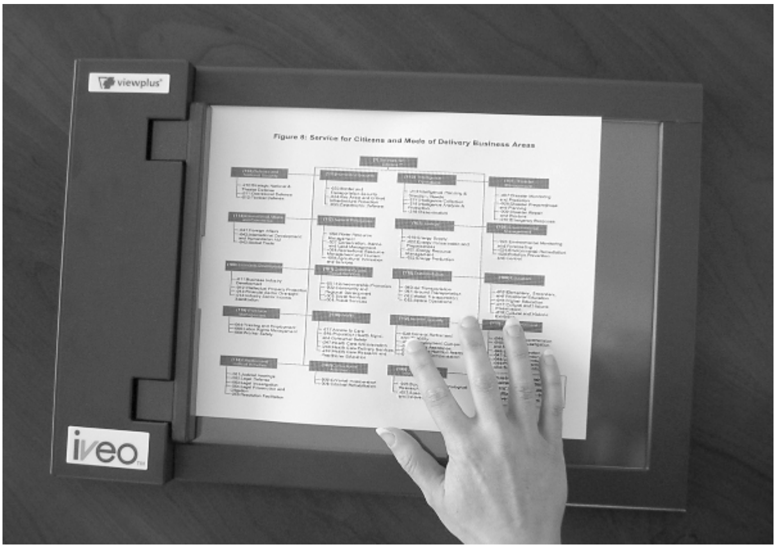


Fig. 3. Photograph of hands examining a print/embossed image made with the new ViewPlus Emprint color haptic printer. The file is a converted PDF image of a document from a US government web site, `http://www.doi.gov/ocio/architecture/fea.htm`. This document is a diagram with a great deal of text and will be much easier to read by people with print disabilities with IVEO™ Release. 2.

3.4 Non-speech Access to Quantitative Data Graphs

Color hue or intensity is occasionally used to represent quantitative information as a function of position in two-dimensional data plots. The most common example is geographic based display of variables such as population density, wealth, presence of natural resources, medical data, etc. Such displays cannot easily be represented accessibly with IVEO™ release 1. Research has recently begun on the feasibility of using audio tone and/or other non-speech audio to substitute for the hue or intensity. The visual variable would be represented by an audio tone whose pitch changes as the user moves the mouse over the diagram. The SVG file could be extended to permit original data to be included so that the audio tone can be derived from data instead of the hue/intensity of the visual image. Introduction of these new display capabilities is expected within one to two years.

Acknowledgements

The authors acknowledge funding by the National Eye Institute, National Institutes of Health for Small Business Innovation Research (SBIR) grants that supported development of IVEO™ hardware and the IVEO™ Viewer. They also acknowledge funding by the National Science Foundation for a SBIR grant that supported development of the IVEO™ Creator application.

References

1. Gardner, J., (1996) Tactile Graphics, an Overview and Resource Guide, Information Technology and Disabilities, 3(4), 1996, `http://www.rit.edu/~easi/itd/itdv03n4/article2.html`
2. Parkes, D. (1988). Nomad: an Audio-Tactile Tool for the Acquisition, Use and Management of Spatially Distributed Information by Partially Sighted and Blind Persons. Proceedings of the Second International Symposium on Maps and Graphics for Visually Handicapped People, King's College, University of London, eds Tatham, A. F., and Dodds, A. G., 24-29
3. Parkes, D. (1991). Nomad: Enabling Access to Graphics and Text Based Information for Blind, Visually Impaired and Other Disability Groups. Conference Proceedings, Vol. 5. World congress on Technology 1991, Arlington, Virginia, 690-714
4. Parkes, D. (1995). Access to Complex Environments for Blind People Multi-Media maps, Plans and Virtual Travel. Proceedings of 17th International Cartographic Conference, Barcelona, V. 2, 2449-2460
5. Parkes, D. (1998). Tactile Audio Tools for Graphicacy and Mobility. A Circle is either a Circle or it is not a Circle. The British Journal of Visual Impairment, 16, 98-104
6. Parkes, D., Brull, M. (1997). It May Not be Easy, but it is Possible a New Form of Literacy. The World Blind, Journal of the World Blind Union, 14, 48-50
7. Loetzsch, J. (1994). Computer-Aided Access to Tactile Graphics for the Blind. Proceedings of the 4th International Conference on Computers Helping People, Vienna 1994, Lecture Notes in Computer Science, 860, Springer-Verlag, 575-581
8. Loetzsch, J., Roedig, G. (1996). Interactive Tactile Media in Training Visually Handicapped People. New Technologies in the Education of the Visually Handicapped, Proceedings of the Conference, Paris, June 10-11, 1996, Burger, D. (Ed) INSERM John Libbey 1996, 155-160
9. Gardner, J., Bargi Rangin, H., Bulatov, V., Kowallik, H., Lundquist, R. (1997). The Problem of Accessing Non-Textual Information on the Web. Proceedings of the 1997 Conference of the W3 Consortium, Santa Clara, CA, April, 1997. `http//dots.physics.orst.edu/publications/www6.html`
10. Bulatov, V., Gardner, J. (1998). Visualization by People without Vision. Proceedings of the Workshop on Content Visualization and Intermediate Representations, Montreal, CA, August 15, 1998, pp 103-108
11. Gardner, J., Bulatov, V. (1998). Non-Visual Access to Non-Textual Information through DotsPlus and Accessible VRML. Computers and Assistive Technology ICCHP'98. Proceedings of the 15th IFIP World Computer Congress, 31 August - 4 September, 1998, Vienna and Budapest, Edwards, A., Arato, A., Zagler, W. eds, Austrian Computer Society on behalf of the IFIP, Vol. 118 pp 136-140

12. Gardner, J., Bulatov, V. (2001). Smart Figures, SVG, and Accessible Web Graphics. Proceedings of the 2001 CSUN International Conference on Technology and Persons with Disabilities, Los Angeles, CA, March 21-24, 2001. `http://www.csun.edu/cod/conf/2001/proceedings/0103gardner.htm`
13. Gardner, J., Bulatov, V. (2004). Complete Access to All Paper and Computer Forms, Tables, Charts. Proceedings of the 2004 CSUN International Conference on Technology and Persons with Disabilities, Los Angeles, CA, March 17-20, 2004. `http://www.csun.edu/cod/conf/2004/ proceedings/162.htm`
14. Walsh, P., and Gardner, J. (2001) TIGER, A NEW AGE OF TACTILE TEXT AND GRAPHICS, Proceedings of the 2001 CSUN International Conference on Technology and Persons with Disabilities, Los Angeles, CA, March 21-24, 2001
15. Gardner, J. (2002). Hands-on Tutorial on Tiger and Win-Triangle. Proceedings of the 2002 CSUN International Conference on Technology and Persons with Disabilities, Los Angeles, CA, March 20-23, 2002. `http://www.csun.edu/cod/conf/2002/proceedings/300.htm`
16. W3C (2005) "Scalable Vector Graphics", retrieved February 2005 from `http://www.w3.org/Graphics/SVG/`

CyARM: Interactive Device for Environment Recognition and Joint Haptic Attention Using Non-visual Modality

Tetsuo Ono[1], Takanori Komatsu[1], Jun-ichi Akita[2], Kiyohide Ito[1], and Makoto Okamoto[1]

[1] Department of Media Architecture, Future University-Hakodate, 116-2 Kamedanakano, Hakodate 041-8655, Japan
{tono, komatsu, itokiyo, maq}@fun.ac.jp
http://www.fun.ac.jp/~tono/
[2] Department of Information and Systems Engineering, Kanazawa University, Kakuma, Kanazawa 920-1192, Japan
akita@is.t.kanazawa-u.ac.jp

Abstract. We have developed CyARM, a new kind of sensing device especially for visually impaired persons, to assist with mobility and detection of nearby objects. This user interface has unique characteristics of giving visually impaired persons the impression of an "imaginary arm" that extends to existing obstacles. CyARM is also a communication device for constructing "joint haptic attention" between impaired and unimpaired persons watching or feeling the same objects. In other words, this device offers a new methodology for judging the others' attentions or intentions without using their eye gaze direction. We verified the efficiency and ability of CyARM for environment recognition through experiments and discuss its possibility as a communication device for realizing joint haptic attention.

1 Introduction

We humans can avoid obstacles while walking, and can stop in our way when blooming flowers catch our attention. Our brain and sensory organs bear a large part of responsibility for controlling such unintentional behaviors (Yoshimoto, 1979). Here, one question arises; "are we humans able to perceive the surrounded environment as it really is?" For example, insects perceive their environment by compound eyes, while bats do by ultrasonic wave generator and receptor. Here it can be said that all animals have developed their own sensory organs and their own ways of environment recognition methods that are suitable for their ways of living style. Therefore, it can be said that they see the same environment differently.

The sensory organs of humans can be roughly divided into the following two groups (Hall, 1976).

- Peripheral receptors: those concerned with the detection of faraway objects, e.g., eyes, ears or noses.
- Contiguous receptors: those used for detection of the nearby surroundings; in other words, sensations received through skins, membranes, or muscles that are tactile sense.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 1251 – 1258, 2006.

However, this categorization is not so strict; for example, sensory stimulation (e.g., radiant heat), is not generally perceived on tactile receptor, however in some cases this receptor can interpret this stimulus. Therefore, there are some possibilities that a person who lost some sensory receptors would compensate these lost functions for using other receptors. Actually, visually impaired persons develop their auditory functions to perceive space in their environment using the relationship between direct and reflective sound waves. However, there might be limits how much the innate sensory organs of humans can compensate the functions of lost sensory organs.

Up to now, although there are not mature technologies to restore lost sensory organs as it were perfectly, it might be possible to suggest new perception mechanism by enhancing the functions of remaining sensory organs. We believe that it is also possible to make non-sensory organs or our physical actions become a kind of new sensory organs. In this paper, we focused on our physical action for compensating the lost sensory organs, especially visual information. Specifically, we developed a new interactive device, called CyARM (abbreviation of Cyber Arm), to provide the visually impaired with a unique and intuitive interface to comprehend their living space. We also discuss the possibilities of CyARM as a communication device for constructing the relationship between visually impaired and unimpaired persons.

2 CyARM

2.1 Motivation

Currently numerous visual aid devices for the visually impaired have been developed instead of white cane, and these devices can be divided into two groups with respect to their user interfaces' modality. The first group is using auditory signals. Specifically, information in surroundings environment is gathered by ultrasonic sensors and transformed into audible sounds that the user can hear (e.g., the high pitch sounds means the obstacle is close to them).The second group is the tactile modality. Information obtained by distance sensors is conveyed to the user through tactile stimuli such as vibrations. Some of these methods are already applied for the commercial products.

Up to now these devices have some critical problems: for example, the devices in the first group transforming "distance information" to "audible sounds" emit the sounds output. However this emitted sound may mask the natural sounds which are indispensable for visually impaired. Because "audio information" is vital for visually impaired to comprehend the spatial condition in their environment. The devices in the second group emit the vibrations with accordance with the detected distance, however it is difficult for users to specify the exact distance from this vibration. They therefore force to create and utilize a kind of mental mapping between "distance" and "frequency of vibration" and this would cause cognitive or mental loads for them.

In order to resolve these issues, we have designed a new sensory aid device for the visually impaired, allowing them to perceive distance and other spatial information without interference to natural sound source and cognitive or mental load (Okamoto et al., 2004).

2.2 Concept

A great deal of researches in sensor technology has been undertaken in recent years. If it succeeded in coupling this sophisticated technology with new interface method, it can create a sensing device that would allow the visually impaired to perceive their surroundings. Since human cannot intuitively recognize artificial signals (e.g., the meanings of vibrations mentioned in the above section), it is essential to propose an intuitive transforming method from the sensory information to certain signals for users that is easy to understand for them.

Suppose that you try walking with your eyes closed, you will attempt to investigate your environment by extending your arms in front of you. This kind of behavior is similar to the functions of insects' antenna or the white cane of visually impaired. Similar to these antenna or canes, when the extend arm touched some objects, one would bent her/his arm at the elbow and stop exploring. On the other hand, there are not object in front of her/him, the arm naturally extends (Fig. 1). Here, it can be considered the bending and extending physical arm motions as a kind of sensory receptor. CyARM was then developed by focusing on this intuitive metaphor of physically arm motion as a sensory receptor. Specifically, it assumed that the users grasps CyARM by their hand and investigates the desired direction with holding this device (Fig. 2).

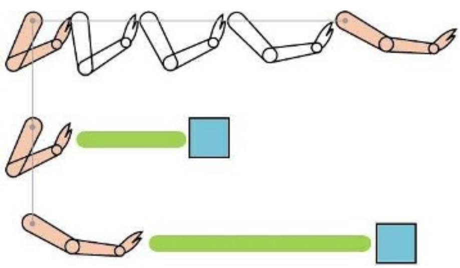


Fig. 1. Metaphor of physical arm motion

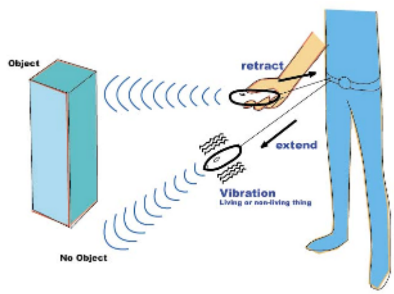


Fig. 2. Concept sketch of CyARM

CyARM is connected to the user by a wire and measures the distance to an obstacle with ultrasonic waves. The tension of the wire is controlled according to the measured distance to the obstacle. If the obstacle is at a short distance, CyARM pull the wire tightly so that the user feels stronger tension and arm was force to bend. This means that the object can be reached just by extending the arm. On the other hand, if the obstacle is far, CyARM stops after giving just enough slack to the wire, so that the user can extend the arm and feels nearly no tension, this means that the obstacle cannot be reached. Therefore, the user can search for obstacles in any direction with holding this device.

2.3 Mechanism

The mechanical structure of CyARM is depicted in Fig. 3. Ultrasonic sensors measure the distance to an obstacle (measurement range is 1.0 – 3.0 m), and the geared motor release or tightens the wire. The wire position is regulated by *P* gain. When the obstacle was found by ultrasonic sensors, the wire is rewound to the initial default position, and its rewinding tension is regulated in accordance with the measured distance. High tension means that the detected obstacle is in a short distance, while low tension means a longer distance. When the users attempt to expand the arm, the device detects a slight displacement of the reel caused by the wire tension, and regulates wire rewinding. The CyARM has the following characteristics.

- Motor: Maxon GP 16 (4.5W) with 29:1 gear head and magnetic rotary encoder.
- Motor Driver: iXs iMDs03-CL
- MPU: Renesas H8/3664
- Ultrasonic sensor's frequency: 38 kHz

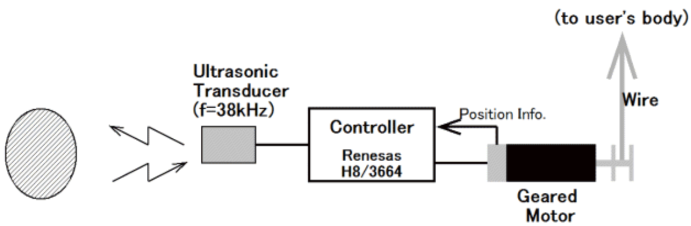


Fig. 3. Mechanical structure of CyARM

Ultrasonic sensors have been placed on the front of its body for users' easy aiming. A hook is attached near the wire release site, allowing for attachment for the user's body (e.g., a belt loop). The prototype CyARM has 500 g weight and its size is 15cm x 10 cm x 10cm.

Fig. 4. Overview of CyARM (prototype)

Fig. 5. Users with holding CyARM in their hands

4 Environment Recognition by CyARM

We carried out two simple testing experiments to clarify the efficiency and ability of the CyARM for environment recognition. The first experiment was to confirm an accuracy of detecting the presence of a stationary object, and the second one is to do an accuracy of detecting the inter-space between objects.

4.1 Experiment 1: Target Presence Detection

The participants of this experiment were four sighted persons with blindfolds and one visually impaired who was completely blind. After given a brief instruction and practice of using CyARM, all participants were asked to stand and hold the CyARM with wearing headphones playing white noise to prevent the external sounds. The task of these participants was to determine whether the static object was present in front of them or not by means of CyARM. As a static object, the whiteboard, one meter width and two meter height, was randomly placed at a distance of about two meters in front of participants. Each participant experienced 20 trials; 10 trials out of 20 were the cases that the object was present, while rests of 10 trials were that it was not.

As a result, the mean percentage of succeeding in detecting the object presence was 90.0% to the object's presence, while the percentage of object absence was 96.0%. The result then suggests that CyARM is efficient device for detecting objects near the user.

Table 1. Result of experiment 1

	recognized as 'exist'	recognized as 'non-exist'
object present	90%	10%
object absent	4%	96%

4.2 Experiment 2: Space Finding Task

We designed this second experiment to examine the following two issues.

1) Is CyARM used for judging whether it is possible to pass through the space between objects?
2) What is the difference between CyARM and commercial counterpart "Tri-sensor" in terms of performance of object detection?

Tri-sensor was developed as a commercial type electric travel aid device for the visually impaired persons. Like bats, Tri-sensor measures the distance between the device and an object by means of ultrasonic waves, and the acquired distance information was converted to audible sound for the user, e.g., high pitched sounds means that there are object close to users.

The participants for this experiment were four sighted with blindfolds. They were asked to find the space between two panels, and then to judge whether it was possible to pass through the space or not. In case that they judged that there was enough space to walk through, they were asked to do so with using holding devices (CyARM or Tri-sensor) (Fig. 6). Each participant experienced 26 trials. 13 trials out of 26 were that participants were asked to use CyARM, while the rest of 13 trials were Tri-sensor. The two panels were set with two different gaps, 0.5 or 1.0 m. If the participants succeeded in passing through 1.0 m gap between panels or in reporting it is impossible to pass trough 0.5 m gap, it was recognized that participants succeeded in this trial.

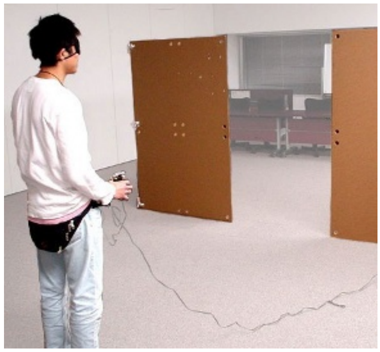


Fig. 6. Finding the space between objects by CyARM

We then calculated percentage of success or failure in the given task. As a result of this experiment, it can be said that both CyARM and Tri-sensor showed high rates of accurate judgment: 92.3% in both devices. This indicates that CyARM could accurately detect the space between objects and enable the user to pass through the space. It also can be said that the usefulness of CyARM for finding gaps between objects is equivalent with the commercialized product like Tri-sensor.

These two preliminary experiments succeeded in showing the usability of CyARM for detecting the static objects or free spaces smoothly. These results also imply that the CyARM allows for safe and accurate orientation and navigation for visually impaired persons.

5 Joint Haptic Attention by CyARM

5.1 Motivation

Joint attention is one of the critical cognitive ability for cognitive development e.g., acquisition of language and Theory of Mind, and is defined as "Watching the certain point or target which the other is watching or pointing." As an acquisition mechanism of joint attention, Baron-Cohen (2001) proposed a schematic diagram. In his diagram, visual information is crucial to establish joint attention. Therefore joint attention is sometime called as "joint visual attention."

However, is joint attention the relationship constructed only by a visual modality? There are some evidences that Shared Attention Module (Baron-Cohen, 2001) can be activated without detecting the eye gaze direction of others. In other words, the joint attention would be realized by other modalities in the same way. Suppose that you would have to touch object, and then feel the other person's hand touching the same object and touching your hand at the same time. In our preliminary research, impaired parents strongly desire to have this sort of shared attentional experience with their children. We aim at building a new communication device which can provide "remote haptic sense" for their creating joint attentional relationship.

5.2 Concept and Mechanism

We are planning to attach a communication device on CyARM for realizing that one person can detect an object that another person is pointing to. This new version of CyARM will offer a new methodology for judging the others' attentions or intentions without using their eye gaze direction. This new CyARM might realize "joint haptic attention" with visually impaired persons (Akita et al., 2005).

Specifically, the new communication device planned to put on the CyARM can be implemented by emitting invisible spot light with a specific frequency modulation and by sensing the other's spot lights by photo detector with demodulator. The spot light and photo detector is set in the same direction. Suppose that there are two users holding this sensing device and the spot light frequency of user A's CyARM is 1 kHz and user B is 0.5 kHz.

If user A and B point the same target like in Fig. 7, User A's photo detector would detect the spot light with frequency 0.5 kHz emitted by User B and User B's one would do the light with 1 kHz emitted by User A (the sensing range of photo detector is like a conical shape). We will implement vibrator for this sensing device; when this device detect the other's emitted spot light, the vibrator would activate and tell its user that someone is pointing the same target. This tactile information would be intuitive for user to know the other is pointing the same target and similar to the actual haptic feeling for users.

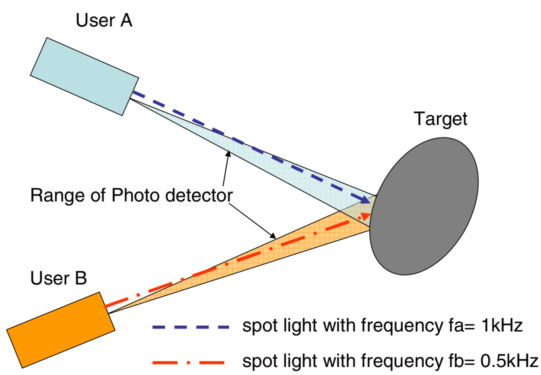


Fig. 7. Concept of sensing method for joint haptic attention

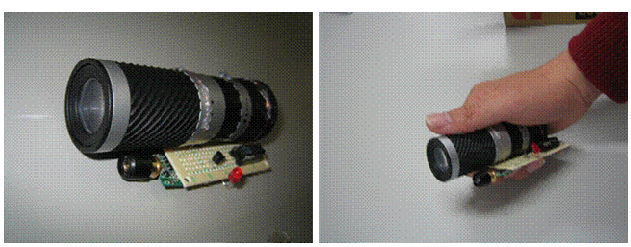


Fig. 8. Prototype sensing system for joint haptic attention

We are now developing the prototype sensing system (Fig. 8), which is separated with CyARM itself. The large cylindrical part is photo detector and the small cylindrical one is spot light emitter. Notice that these two parts is fixed in the same direction, so that the pointing and searching direction is the same and these both behaviors were done simultaneously.

6 Conclusions

We have developed CyARM, a new kind of sensing device especially for visually impaired persons, to assist with mobility and detection of nearby objects. A wire is connected to user's body and controlled by CyARM according to the measured distances to obstacles. This user interface has unique characteristics of giving visually impaired persons the impression of an "imaginary arm" that extends to existing obstacles. CyARM is also a communication device for constructing "joint haptic attention" between impaired and unimpaired persons watching or feeling the same objects. We verified the efficiency and ability of CyARM for environment recognition through experiments and discussed its possibility as a communication device.

References

1. Akita, J., Ito, K., Komatsu, T., Ono, T. and Okamoto, M.: CyARM: Direct Perception Device by Dynamic Touch, In Proceedings of International Conference on Perception and Action (ICPA13), (2005)
2. Baron-Cohen, S.: Mindblindness, MIT Press, (2001)
3. Hall, E. T.: The Hidden Dimension, Peter Smith Publisher Inc., (1976)
4. Okamoto, M., Akita, J., Ito, K., Ono, T. and Takagi, T.: CyARM: Interactive Device for Environment Recognition using a Non-Visual Modality, In Proceedings of International Conference on Computers Helping People with Special Needs (ICCHP2004), pp.462-467, (2004)
5. Yoshimoto, T.: Here with Finger, Hokkaido University Press (in Japanese), (1979)

Marking in the Surroundings by Data-Carriers for the Visually Impaired

Hisayuki Tatsumi[1], Yasuyuki Murai[2],
Nobuyuki Nagai[1], and Masahiro Miyakawa[1]

[1] Department of Computer Science, Tsukuba University of Technology (NTUT)
4-12-7 Kasuga, Tsukuba, Ibaraki 305-8521, Japan
{tatsumi, nagai, mamiyaka}@cs.k.tsukuba-tech.ac.jp
[2] School of Pharmacy, Nihon Pharmaceutical University
10281 Komuro, Ina-cho, Kita-Adachi-gun, Saitama 362-0806, Japan
murai@nichiyaku.ac.jp

Abstract. We are conducting a research project aimed at building an information secured environment for the visually impaired, especially in our campus, by combining data-carrier technology and networks. We use RFID tags and networks to keep the environmental information and to trace the changes in the environment. Our approach is the following: 1. In the mental map kept by the visually impaired person landmarks play a significant role, 2. We can turn any object (location) in the surroundings into a landmark by attaching an RFID tag to it, 3. The changes in the environment could be recognized, at least roughly, by recognizing changes of landmarks. We outline two methods to trace the changes of locations of RFID tags in the environment. The first one is to detect the moves of active RFID tags directly by analyzing the signals emitted from them. The second one is to trace the moves of passive RFID-tags by autonomous cruising robots. In both cases the changes are kept and updated in the network, and then conveyed to a visually impaired person according to his inquiry.

1 Introduction

In the environment which is dynamic and subject to spontaneous changes, without the presence of an interpreter, the knowledge of the visually impaired person about the environment is kept old and so he often is not aware of the changes in it. The authors are trying to build an information ensured environment for the visually impaired students in the college life (all students are visually impaired though its degrees vary large). For this purpose, we need to know the changes of the environment real-time, or at least with a small delay. We show that data-carrier technology together with networks can provide the visually impaired with an environment in which they can get information about the surroundings that are subject to change. We describe ongoing project of marking the environment using RFID tags and maintaining its dynamic changes by RFID sensing robots. All these can be viewed as providing environmental information for the visually impaired.

Take the walking of a visually impaired person as an example. He obtains his walking information through the vibrations of his white cane which he is manipulating along tactile blocks on the pathway. He becomes able to walk through the

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 1259 – 1264, 2006.

pathway only after repeated trials during which time he remembers every clue on the pathway such as the edges of guardrails and corner stones. Based on this observation we notice two important things for the aid of the visually impaired: one is marking in the environment and the other is keeping track of the moves of the marks. We give two solutions for keeping track of the moves of the marks in the room: one is by analyzing a signal (direction and strength) emitted by active RFID tags attached to the object, and the other is by using autonomous RFID sensing robots which cruise and check the moves of the object by reading the RFID tag.

The material we report here is largely borrowed from our almost simultaneous report in [1]. However, in this paper we tried to incorporate directions of our future research plan for the aid of the visually impaired.

2 Marking in the Environment by RFID

2.1 RFID Tag

There are two types of RFID tags. The first type, passive RFID has no electric power supply. Its electric power to drive memory circuit is supplied remotely from the reader by electro-magnetic induction, so its memory capacity is limited (few 100-bytes). Also the communication range is short (from few centimeters to few 10-centimeters). These restrict its use. But it allows communication in both directions (i.e., writing in the tag is allowed as well as reading). The second type, active RFID has its own power supply (battery) in it. Its communication range is much larger (from several meters to several ten meters). Because of long range communication it usually allows only the emission of signals. By analyzing a signal received from the active RFID tag we may know the direction and the distance of the RFID (object) as it serves as a small-powered radio beacon.

2.2 Marking

We list several situations where we need marking for visually impaired students.

(1) Identifications of locations in the campus
(2) Walking guide in the school buildings
(3) Switching of the walking modes at the start and at the end of stairways
(4) Guidance of placement of personal belongings (e.g., books)
(5) Moves of objects (e.g., desks) in the classroom or in the hallway

We focus on Situation (5). It arises in a laboratory or in a hallway where environmental change is caused by the moves of objects; without an adequate aid it is very difficult for the visually impaired to understand the new situation and take a suitable action against the change. To solve this situation we mark every object in the room (the environment) by attaching an RFID tag to it and register its property (location etc.) in a database on the network. To cope with the spontaneous moves of the objects we track them by a cruising robot which can search and sense the RFID tag and update its location in the database. This is described in Section 4. In the next section we describe shortly about a possible use of active RFID tags to detect moves of the objects directly by analyzing signals emitted from the RFID tags.

3 Active RFID Tags and the Detection of Moves

In Fig.1 we indicate active RFID devices used in our experiment. The tag emits signals (its identification number consisting of 6 alphabets and 8 digits) in every 7 seconds. Also a vibration sensor is set in the tag to detect the move of the tag and it keeps emitting signals (its identification number prefixed by the symbol M) twice a second while vibration is being sensed. So sensing the start and the end of the move of the object is easy. If we know differences both in directions and in intensities of the signals before and after the move, we may calculate the new location from the old one. In order to see this possibility we have done a preliminary experiment detecting the intensities of signals as a function of distance (and possibly of directions) from the receiver. As a result we have seen that it is hardly possible to distinguish two tags by its intensities set apart by 20 cm placed in the extremely short distance of 40 cm from the receiver [1]; we need a more careful study on this theme. However, active RFID has a promising future use for the visually impaired.

(a) A tag (b) A reader (receiver)

Fig. 1. Active RFID devices

4 Tracking the Moves of Marks by RFID Sensing Robots

In this section we describe a scheme in which the changes of the environment can be known by reading the passive RFID tags deployed in the environment (with the help of network the tag can contain sufficient information). If a visually impaired person is not able to get an access to RFID tags, instead he can ask a server thorough the network about the change, thereby avoiding the help of other (sighted) people.

Our goal is a question-answering system about environmental information. When a visually impaired person senses a change in the environment, he asks the robot for the detail. This includes the situation of avoidance of the obstacles caused by a temporal move of the object, say, in the hallway. Our system should give adequate answers to the following questions: “Where am I ?”; “In what situation am I ?” For this purpose the RFID sensing robot detects the changes of environment through autonomous cruising and keeps the environmental knowledge always up to date. Here arises a navigation problem of the robot. We briefly describe our navigation program (a prototype) in the following paragraph together with the hardware.

Fig.2 shows our sensing robot from a front view. The robot is built on the platform of ER-1 (a kit by Evolution Robotics Inc. [2]) attached a high-performance passive RFID reader/writer with 25cm effective antenna range. This reader/writer and a passive RFID communicate with 13.56 MHz electro-magnetic induction wave. The robot is autonomous. The software named RCC (Robot Control Center) installed in the laptop carried on ER-1 controls both the sensor system consisting of CCD camera and IR (Infrared Radiation) sensors and the driving system consisting of three wheels which are driven by batteries.

The navigation program is sent to RCC via telnet channel. This makes the robot controllable in two modes: one is manipulating RCC via telnet from a remote computer and the other is making the robot run autonomously by loading the navigation programs in the laptop as well as RCC and telneting the navigation program to RCC within the laptop. Environmental information obtained by the robot (the content of RFID tags and the location of the robot) is sent to a server and kept there for a possible inquiry. In this way environmental knowledge is maintained in the network.

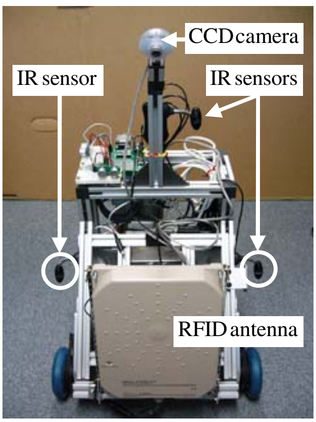


Fig. 2. Our passive RFID robot

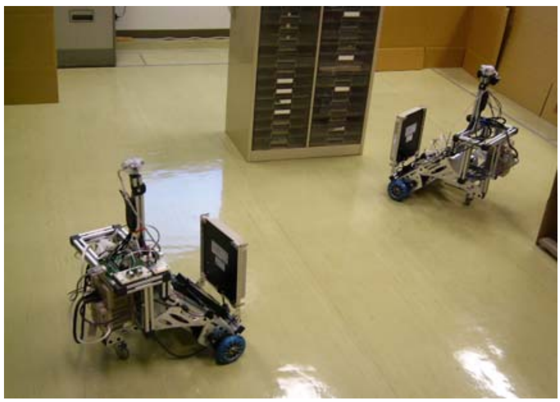


Fig. 3. Environment in our experiment

Experiments: We have assumed the following simple situation as our environment: in a rectangle-shaped room (5m x 3m) there are several obstacles and objects (all are rectangle-shaped) scattered in a fairly sparse way (see Fig.3). The task of our RFID sensing robots is to make a cruise in such a way that it avoids obstacles and searches objects (to which RFID tags are attached) as many as possible (hopefully all) in the room. We have implemented an intuitive navigation program based on the IR sensor information (currently no visual information from CCD camera is used). The program consists of two routines: one is a searching program (the robot searches the objects) and the other is an approaching program which controls approaching of the robot to a specified object. A detail of our planned navigation principle is given in [3,4]. Fig.4 shows the robot reading a passive RFID tag attached to an object. Our navigation program has successfully avoided both collisions and obstacles and found almost all objects.

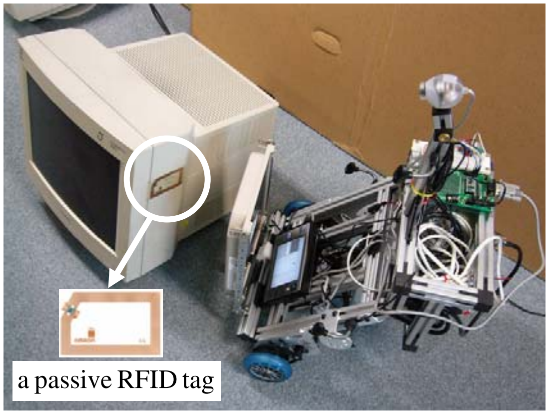


Fig. 4. RFID robot reading the tag

5 Future Research Directions and Discussions

Currently we are designing a program which constructs a database of the environmental changes. Our goal is the construction of such environmental database that an efficient navigation of the RFID sensing robots maintains the environmental information in the database. To make a map of the environment, as a coming next step we use a high-performance autonomous robot Scorpion (See Fig.5, a product from Evolution Robotics Inc.). This robot is controlled by a program made by using ERSP (Evolution Robotics Software Platform) which consists of several hundred thousand lines of C++ codes. One of the most important technologies of the autonomous robot is the ability to automatically build a map of the environment and localize the robot within that map. The ERSP's navigation components include the powerful visual Simultaneous Localization and Mapping algorithm (vSLAM). Fig.6 shows the grid map of our room, which is obtained by using vSLAM.

Fig. 5. Scorpion robot

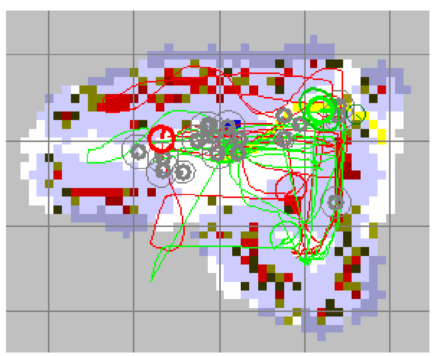


Fig. 6. A grid map of our room

The purpose of our project is to build an information ensured environment for visually impaired students in the college life. Keeping in mind the importance of the roles of marks in the environment and of keeping track of moves of the marks for the visually impaired, we have proposed an acquisition scheme of environmental information which is subject to change. Namely, we have proposed the use of RFID tags in marking in the environment and the RFID sensing robots to trace the environmental changes. A possible use of active RFID tags for the detection of moves of objects is also mentioned. Data carrier technology is expected to bring benefit, especially for the visually impaired.

Acknowledgment

This research was done under NTUT (National Univ. Corp., Tsukuba Univ. of Tech.) Education Promotion Grant 40: Aiding environmental information acquisition for the visually impaired by RFID and sensor-network.

References

1. Tatsumi, H., Murai, Y., Miyakawa, M., Tokumasu, S.: Marking in space and acquisition of environmental information by data-carrier for the visually impaired, WAC (World Automation Congress) 2006 Proc. 5th Int. Forum on Multimedia and Image Processing, Budapest, Hungary (July 2006), to appear
2. http://www.evolution.com/er1/
3. Tatsumi, H., Murai Y., Miyakawa M., Tokumasu S.: Providing environmental information by data-carrier + networks, IEEE Proc. 2005 Int. Conf. Systems, Man and Cybernetics, Hawaii, USA (2005) 3628-3633
4. Murai Y., Tatsumi, H., Tokumasu, S., Miyakawa, M.: Autonomous navigation of RFID-sensing robots, IEEE Proc. 2005 Int. Conf. Systems, Man and Cybernetics, Hawaii, USA (2005) 3634-3639

A Study and Development of the Auditory Route Map Providing System for the Visually Impaired

Takafumi Ienaga[1], Michito Matsumoto[1,⋆], Masahiro Shibata[1,⋆⋆], Nobuyuki Toyoda[2], Youko Kimura[2], Hiroshi Gotoh[2], and Takehito Yasukouchi[2]

[1] Institute of Systems and Information Technology / KYUSHU, Fukuoka SRP Center Building 7F, 2-1-22, Momochihama, Sawara-ku, Fukuoka, Japan
ienaga@isit.or.jp, michito@pu-toyama.ac.jp
[2] Fukuoka Welfare Center for Persons with Disabilities

Abstract. These days, we are able to get updated visual map freely from web sites. However, auditory maps for the visually impaired were not able to get easily. So, we studied about the system which enables us to generate and provide the auditory map via Internet. In this paper, we described about the production rules of our auditory map. And, we also described about the questionnaire and two experiments to determine the information which was put into the map. The results of experiment suggested that it is important whether the subjects were able to find the point for changing their directions to reach to the destination. And we applied that suggestion to our prototype system. And, finally, we also described about the problems of our system and our future works.

1 Introduction

We use a map in order to travel in a place which is unfamiliar to us. These days, we are able to get a visual map easily on the web sites (e.g. Google Maps [1]), and use such a updated map. Against this situation, some travel aid systems for the visually impaired were studied and developed (e.g. MoBIC [2]). However, these systems had problems about a covering area, the newness of the information, the dependency of the road condition, and so on. Therefore, especially in Japan, we are not able to easily get updated and useful auditory or tactile maps for the visually impaired. So, we studied about the auditory route map which suited road conditions of Japan. And, we also studied about the system which produced an auditory route map automatically and provided the map via the Internet. Here, there are some systems which guide the visually impaired to the destination with measuring the user's current position (e.g. Trekker [3]). In these systems, the precise measurement of the user's current position is required. However, it is impossible to measure precisely with GPS in the downtown which has many high buildings and indoors. And, there is a limitation in a covering area, to use

⋆ He moved to Toyama Prefectural University.
⋆⋆ He moved to Kyushu University Venture Business Laboratory.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 1265–1272, 2006.

RFID tags for the purpose of position measurement (e.g. [4]). Therefore, in this study, we refered to only pre-planing using the auditory map. And, we assumed that the users of our system were the visually impaired who were able to walk independently using a white cane or a guide dog.

2 A Study of Auditory Route Map

2.1 The Typical Senario for Using Our System

The scenario of using our system is shown as below. In the 1st step, a user accesses the system and requests an auditory route map by sending some information, such as a departure place, a destination place, and the information about the user him/herself. In the 2nd step, the system produces an auditory route map based on the geographic information, the additional environmental information, and the information about the user. Here, the additional environmental information means that the cues to be used by the visually impaired for walking and orientating purposes. And then, the map is produced by the system according to the userfs characteristics. In the 3rd step, the user gets the auditory route map and records it into his or her audio device. In the 4th step, the user brings the audio device and listens to that audio route map while walking. As these steps, the user is able to reach the destination place by him/herself.

2.2 The Basic Production Rules of Instructions for Auditory Route Maps

First of all, we determined the basic production rules. In order to determining the rule, we referred to the other Japanese auditory maps [5,6], and considered the actual trainings for acquiring the skills to walk independently at Fukuoka Welfare Center for Persons with Disabilities.

At first, we created the pedestrian networks. Here, if the width of the road was less than 4 meters, we didn't distinguish the sides of the road. And, we separated this network into some segments by entrances of crosswalks or cross points. Here, if the segment consists of only a crosswalk, we merged it into the previous segment. And then, we generated the auditory map shown as below.

Our auditory route map was separated into two parts, the summary part and the detail part. In the summary part, the map indicates (1) the departure place and the destination place as the whole route, (2) the current position of the user, (3) the direction to walk along and the methods for orientating, (4) the name of the destination, (5) the direction to the destination, (6) the time for walking until getting to the destination, (7) the number of sections in the detail part, (8) the features of the route, and (9) the fact that the chime rings at the start and end of each section. And, in the detail part, the instruction was divided into some sections. And, each section were correspondent to each segment. The instruction of each section describes (1) the departure and the destination place of the section, (2) the user's position and direction, (3) the instructions about the road to walk along, (4) the instructions about the environments around the road, (5) the instructions about the destination place, and (6) the instructions about the direction to the destination place of the section.

2.3 The Questionnaires for Specification Decision

In order to determine what information should be put into the auditory route map, we did questionnaires targeting 50 persons who were visually impaired. They lived in around Fukuoka City in Japan, and almost all of them were able to walk independently around their house. The question was "When you walk independently with a white cane (or a guide dog), do you use this information or cue?". Here, the information or cues is shown as Table 1. At first, we selected the information which was answered "Yes" by more than 35 persons (see Table 1). And, considering the difficulties of collecting data and the educational viewpoint of the skills for independent walk, we selected the information. That was "Route", "Direction", "With or without sidewalk", "State of cross points", "Traffic signals", "Environment around the roads", "Condition of sidewalk", "Length of the road", "Traffic signal with sound device", "Textured paving blocks", "Surface of road", "Signs on the road", "Slope", and "Width of the road".

Table 1. The result of the questionnaire against 50 persons. The frequency means the number of persons who answered "Yes"

Classification	Information or Cues	Frequency	Classification	Information or Cues	Frequency
General	Route	45	Tactile	Textured paving blocks	47
	Direction	41		Surface of road	40
	With or without sidewalk	37		Signs on the road	35
	State of cross points	42	Olfactory	Paticular smell	40
	Traffic signals	43	Visual	White line	6
	Environment around the roads	35		Textured paving blocks	8
	Condition of sidewalk	37		Street lights	14
	Width of the road	29	Others	Slope	38
	Length of the road	37		Wind	29
Auditory	Sound of cars	41		Sun	29
	Particular sound	39			
	Traffic signals with sound device	47			

2.4 Experiment 1

We created the auditory map based on the production rules and the information shown as above. And, we did the experiment using this created map. The purposes of this experiment were to clarify whether the visually impaired were able to get to the destination with independently walk using this created map and to re-extract the information according to the results of the experiment.

The subjects were 13 persons, who were the visually impaired. The route was set up in the downtown of Fukuoka City (Fig. 1). And, the roads were

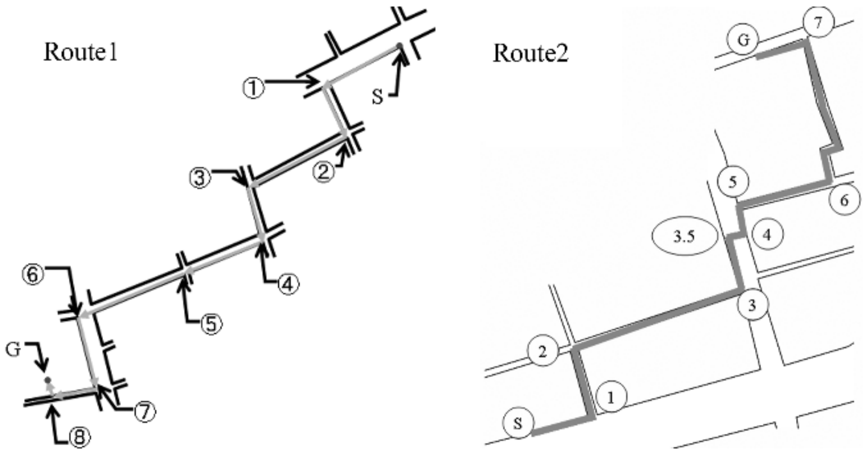


Fig. 1. The routes of the experiments. Route 1 means the route of Experiment 1, and Route 2 means the route of Experiment 2. Here, "S" means the departure place, and "G" means the destination.

sidewalks, mainly. All subjects had no experience to walk along this route. The experimenter was able to get to the destination by walking along this route in about 20 minutes. Before the experiment, the subjects listened to the auditory map in several times, and grasped the whole route roughly. After the starting signal from the experimenter, the subjects started to walk independently. Here, the subjects were able to listen to the auditory map as much as they wanted in anytime. When the subject reached to the destination or 45 minutes passed after the starting signal, the experiment finished. And, when the subject said "I will give up." or the subject walk far from the correct route by 2 blocks, the experiment also finished. After the experiment, we did the questionnaire. The 1st question was "Please choose five required and five non-required information for you, when you walk independently.". Here, the alternatives were shown in Table 3. And, 2nd question was "Please choose four required environmental information and four non-required environmental information for you, when you walk independently". Here, the alternatives were shown in Table 4.

The result of the experiment is shown in Table 2. 10 of the 13 subjects reached to the entrance of the destination, which is the check point "8". Considering this fact, our auditory map seemed to basically useful to walk on the sidewalks. How-

Table 2. The relationship between the last check point where subjects reached to and the number of the subjects

Last check point	The number of subjects	Last check point	The number of subjects
G	5	2	1
8	5	1	1
4	1		

ever, 5 of the above 10 subjects were not able to get to the destination. On this route, the road between the check point 7 and check point 8 didn't have any sidewalks, and there were no signs at the check point 8. So, they were hard to find the check point 8. Table 3 and Table 4 shows the result of the questionnaire. Considering the result of the questionnaire, the subjects felt that the information shown as below were non-required, that was, “Width of the road”, “Position of the textured paving blocks on the road”, “Traffic signal's sound type”, “Position of the sound device”, “Material of the road”, “Fell of the road”, and “Width of crosswalk”. And more, the subjects felt that the environmental information shown as below were non-required, that is “Vehicle's road along the sidewalk”, “Utility pole”, “Automatic vending machine”, “Street sign”, and “Street light”.

Table 3. The result of the question1. Here, one of the 13 subjects answered no “Non-require” Information.

Information	Require	Non-require
Summary	3	1
Current direction	9	0
With or without sidewalk	9	0
Width of the road	0	7
Length of the road	5	3
With or without textured paving blocks	7	1
Position of textured paving blocks on the road	0	6
With or without traffic signal	11	0
With or without sound device	6	0
Traffic signal's sound type	0	5
Position of the sound device	0	6
Material of the road	1	8
Feel of the road	0	8
Shape of crosspoint	1	2
With or without textured paving block for warning	7	2
Width of crosswalk	1	10
Length of crosswalk	4	0
Environment around the road	1	1

Table 4. The result of the 2nd question. Here, one of the 13 subjects answered 5 “Necessary” Information.

Information	Require	Non-require	Information	Require	Non-require
Guard rail	7	0	Buildings	7	2
Parking lot	6	2	Bus stop	3	0
Border	5	2	Entrance of the subway	2	3
Vehicle's road along the sidewalk	2	5	Obstacles on the road	3	2
Utility pole	0	3	Automatic vending machine	2	9
Slope	9	1	Street sign	0	12
Shops	7	0	Street lignt	0	11

In consideration of these result, we did re-extraction of the information which were put into the auditory map (Here, we didn't remove the information about "Width of the road", because of the educational viewpoint of the skills for independent walk). And, we designed the experiment on the different route.

2.5 Experiment 2

We created the auditory map based on the production rules and the above re-extracted information. And, we did the experiment using this created map. The purposes of this experiment were to clarify the problems of these maps. The subjects were 12 persons, who were the visually impaired. The route was set up in the downtown of Fukuoka City (Fig. 1). All subjects had no experience to walk along this route. The experimenter was able to walk along this route in about 20 minutes.

Before the experiment, the subjects listened to the auditory map in several times, and grasped the whole route roughly. After the starting signal from the experimenter, the subjects started to walk independently. Here, the subjects were able to listen to the auditory map as much as they wanted in anytime. When the subject reached to the destination or 45 minutes passed after the start, the experiment finished. And, if the subjects lost their way, they were able to use the tele-support [7]. Here, the tele-support meant that the supporter who was at remote place instructed the subjects how to get to the nearest check point and how to get their orientation.

The result of this experiment showed in Table 5. And, the subjects used the tele-support the most around the point 6. In this experiment, the route consisted of the paths without sidewalks mainly. And, the road between the check point 6 and the check point 7 was narrower than the road between the check point 5 and the check point 6. And, there were no signs on the road at the check point 6. So, it seemed that the subjects were hard to find the check point 6.

Table 5. The relationship between the last check point where the subjects went through and the number of the tele-support. Here, "6.5" means in the crank which placed between "6" and "7" on Route 2.

Check point	The number of the tele-support	Check point	The number of the tele-support
1	1	6	6
2	2	6.5	3
3	1	7	2
3.5	3	G	1
5	2		

2.6 Total Discussion

The results of above experiments suggested as follows: the result as to whether the participant was able to reach the destination depended on the result as to whether he or she was able to find the points for turn like the cross points and the entrances of the crosswalk. That is, the route should be selected and the instructions should be produced according to the user's ability of finding the point for turn, not only the difficulties of roads to walk along.

3 Development of the Prototype System

We developed a prototype system of the auditory route map providing system for the visually impaired based on the above results. That is, the auditory route map was produced based on the basic production rules shown as above, and its information was selected based on the result of the above questionnaires and experiments. The system worked on Linux platform, and it was written using PHP. And, the system required the software shown as below; apache, mapserver, postgresql, postgis, and so on.

Another feature of our system was that the nodes for pedestrian's networks had its weight for finding the most appropriate path to take. It was designed based on the results of the above experiment. That is, finding the points for turn was the most important for a user to reach the destination successfully. In general, in order to solve the shortest path problem with Dijkstra's algorithm etc., only the edges for pedestrian's networks have the weight. So, in the case of solving the shortest path problem, the node's weight was added to the edge's weight temporarily. By applying this procedure, the system enabled the user to find the most adequate path to walk along. Besides, in order to make the system work according to individual characteristics, the system had the personal data table which consisted of the type of edge and the type of nodes and their weights. This design enabled to produce the personal maps according to the individual characteristics.

Now, we developed just prototype system which had above mechanism. So, we have not evaluated that the system was able to select adequate route for the individual characteristic, yet. And, this system had still many limitations, for example, the system was able to provide the maps only as text datum. We will improve this system in our future works.

4 Conclusion and Future Works

In this paper, we described about the auditory route map providing system, especially its basic concept, the questionnaires and experiments for the specification determination, and its prototype system. In our future works, we will be able to provide the maps as audio datum, and improve its user interface for the visually impaired. And, we will also study how to collect the information which was put into the auditory route map efficiently and how to weight the nodes and edges adequately.

Acknowledgement

A part of this research was supported by the Japan Society for the Promotion of Science, Grant-in-Aid for Scientific Research (B(2)-16300192).

References

1. Google Maps. <http://maps.google.com/>, (last access 2006-01-20).
2. T. Strothotte, H. Petrie, V.Johnson, and L. Reichert: MoBIC: user needs and preliminary design for mobility aid for blind and elderly travellers. The European Context for Assistive Technology, pp. 348-352, (1995).
3. Adaptive Technology Consulting – A GPS system for the Blind. <http://www.adaptivetech.net/products/trekker.htm>, (last access 2006-04-15).
4. S. Willis and S. Helal: RFID Information Grid for Blind Navigation and Wayfinding. The Proceedings of the Ninth IEEE International Symposium on Wearable Computers (ISWC 2005), pp. 34-37, (2005).
5. H. Eguchi: The Study and Usage of Auditory Maps in Case of the Walk Training for the Visually Impaired. The History of Fukuoka National Rehabilitation Center for Persons with Visually Impaired in these 10 Years, pp.171-179, (1979). (in Japanese).
6. The Route Guidance. <http://www.kotonavi.jp/index.html>, (last access 2005-04-20). (in Japanese).
7. T.Ienaga, M.Matsumoto, N.Toyoda, Y.Kimura, H.Gotoh, M.Shibata, T.Yasukouchi: Travel Aid System with Auditory-map and Video Phone for the Visually Impaired. The 21st Annual International Technology and Persons with Disabilities Conference, (2006.3).

Identification of Acoustic Factors for Perception of Crossability Common to Blind and Sighted Pedestrians

Takayuki Shiose[1], Kiyohide Ito[2], and Kazuhiko Mamada[3]

[1] Kyoto University, Sakyo-ku, Yoshida-Honmachi, Kyoto, Japan
[2] Future University-Hakodate, 116-2 Kameda-Nakanocho, Hakodate, Hokkaido, Japan
[3] School for the Visually ImpairedCUniversity of Tsukuba, 3-27-6 Mejirodai,Bunkyo-ku,Tokyo 112-8684, Japan

Abstract. This paper describes which auditory information affects the accuracy of "perception of crossability" for blind people. We have created a "virtual 3D acoustic environment" in which subjects feel a car passing in front of them to help them cross a road safely. The system is theoretically based on acoustic "time-to-contact" information, which is the most important concept in Ecological Psychology. Experimental results suggest that blind people tend to estimate the acoustic "time-to-contact" significantly longer than did sighted people. However, there are no significant differences between their sensitivities toward the speed of a moving sound source and the gain level of indirect sounds.

1 Introduction

We have developed a "virtual 3D acoustic environment" to help pedestrians with severe visual impairment train to cross roads safely, in which they can hear moving sound sources as if vehicles were passing in front of them [6]. It is difficult to accurately specify which acoustic features affect a pedestrian's level of "crossability." "Crossability" is a sense of perceiving the capability that a person can cross a road without being hit by a vehicle.

In our previous work [6], we demonstrated that it is possible to clearly distinguish subjects who can understand road situations accurately and those who cannot by focusing on individual variations of the acoustic clues they use. Those results led us to focus on accurate perception of speed for accurate understanding of the traffic situation. In this work, we examine whether subjects can perceive speed accurately by using acoustic information only and determine how responsive they are to indirect sound. Experimental results exhibit the same trend for a variety of direct and indirect sounds in cases of both sighted and blind people. Based on these results, we propose a new experimental method in consideration of the individuality of acoustic clues. Finally, we propose a guideline for designing a training menu.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 1273–1279, 2006.

2 Research and Methodological Approach

The purpose of the present study is to find what acoustic skills are necessary for blind people to perceive the crossability of a road. To cross a road safely, one needs to accurately estimate the following two time types:

1. Time remaining until a vehicle arrives at the pedestrians position (time-to-arrival t_a)
2. Time required when crossing a road, estimated from the width of the road (time-to-cross t_c)

If pedestrians wrongly estimate the ta and tc, they may collide with a vehicle. We stand on the viewpoint of ecological psychology, which is known to be a useful study for understanding the time-to-arrival estimation process.

Figure 1 shows the experimental setting. First, each subject was instructed to localize where the white cane made knocking noises, after which the sound source was set to pass about 2 m in front of the subject at a certain speed, which was randomly changed from 10 to 40 km/h. The subjects were instructed to listen to the acoustic information and to click twice, clicking first when they noticed the car sound passing over the white cane knocking point, and second when they heard the sound of the passing car right in front of them. Note that sighted subjects were instructed not to watch the screen.

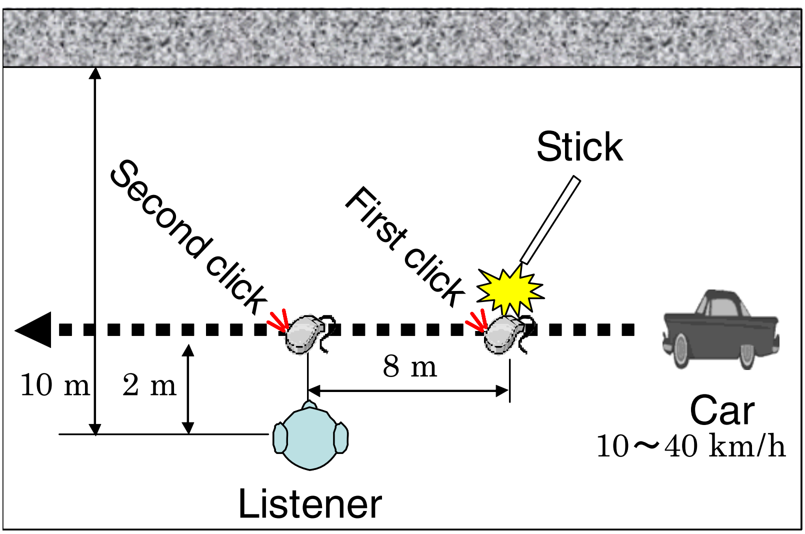


Fig. 1. Top view of this experimental setting

First, the subjects were instructed to locate the sound source moving at different speeds (from 10 km/h to 40 km/h at 5-km/h intervals, presented randomly) in 40 trials. Next, they were instructed to locate the sound source moving at a constant speed of 25 km/h under different indirect sound conditions (eight conditions from -0d B to -60 dB in 10-dB increments and no indirect sound) in 40 trials. We only told the subjects that there were two sets of 40 trials under different conditions, and they were not informed about the details of those conditions. The distance from the point of the knocking noises to the front of a subject was set to 8 m.

3 System for Creating Acoustic Environments

Figure 2 shows the architecture of our system, which consisted mainly of three audio recorders (Roland AR-3000); two 3D sound-space processors (Roland RSS-10); and a headphone system (SONY MDR-DS8000). A moving sound source and white cane knocking noises were transmitted from each "audio recorder" to each sound space processor, which, in a virtual 3D sound space, can simulate acoustic environments including reflected sounds, reverberations, and the Doppler effect.

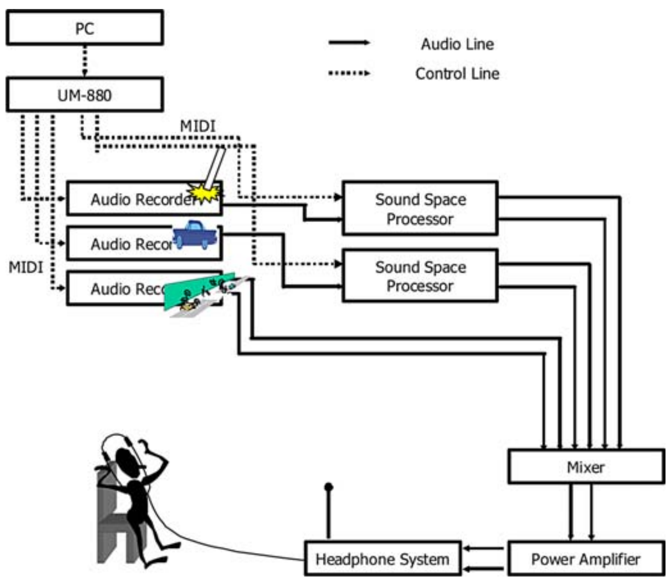


Fig. 2. System architecture for this experiment

The RSS-10s each have four sub-processors: a binaural processor, a distance processor, a reverberation processor, and a delay mode. These sub-processors are based on Head Related Transfer Functions (HRTFs), which are measured in an anechoic chamber. This enables subjects to not only localize stationary sound sources but to also detect the direction of a sound. The RSS-10s were controlled with a PC (IBM IntelliStation Z Pro) through a MIDI interface (Roland UM-880), and the PC output generated sound to the headphones through a mixer

(Roland VM-7200 and VM-C7100). The experimenter can also design many different sound image trajectories using exclusive software (Roland's Creating 3D Sound and Microsoft's Visual C). Finally, ambient sounds, which are assumed to occur in real road-crossing situations, were set into the virtual acoustic environment by the third audio recorder.

4 Experimental Results and Discussions

In this work, we examined whether subjects can accurately perceive speed by using only acoustic information and their sensitivity to indirect sound. Thirteen sighted subjects and 13 blind subjects, aged between 21 and 50 years old, participated in the experiment. We focused on how correctly subjects estimated the 8-m distance from a point where knocking noises were made to the front of a subject. This distance was constant in this experiment. Therefore, we can say that a subject has accurate perception of speed if he or she thinks a car comes early under high car speed conditions and it comes late under low car speed conditions.

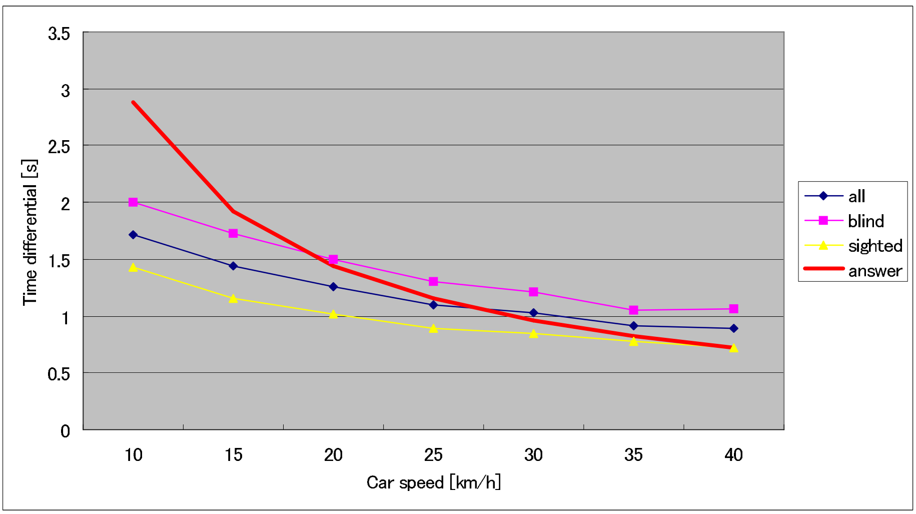


Fig. 3. Accuracy of locating moving sound source vs. change of moving speed

Figure 3 shows the time differences between two clicks under different car speed conditions. The differences between the correct answers and answers given by subjects show how they can estimate the distance correctly. The correct answer varies with car speed as described by the red line.

We used a two-way analysis of variance (ANOVA) to compare means among groups. *Post hoc* analyses were performed with the Bonferroni/Dunn procedure when the F ratio for the ANOVA was significant at $p < 0.05$. In the following table, VI denotes "Visual Impairment" and Sp means "Moving Speed of Sound Source."

Table 1. Two-way ANOVA table for changing parameters of moving speed

Source	Seq. SS	df	MS	F	P	
	71.32794	24	2.971997	60.84272	0.0000	**
VI	8.39369	1	8.393696	2.824261	0.1058	
Speed	14.20917	6	2.368195	48.48169	0.0000	**
VI*Sp	0.53274	6	0.088791	1.81773	0.0996	
Residual	7.03399	144	0.048847			
Total	101.49755	181				

Note: * $p < 0.05$, ** $p < 0.01$

The main effect of the factor "Speed" was significant (p<0.01). On the other hand, the main effect of the factor "Visual Impairment" was not significant (p=0.1058). The Bonferroni/Dunn test revealed the trend that there were significant differences among different speed conditions at low speed (around 10km/h) and no significant difference at high speed (around 40km/h). This indicates us that both blind and sighted people tend to underestimate to a greater extent how long a car takes to arrive at the spot the faster it moves. There was no significant interaction effect, here, although only in cases of blind people, differences between estimated times for different car speed conditions becomes smaller if the car moves faster.

We used a two-way analysis of variance (ANOVA) to compare means among groups. *Post hoc* analyses were performed with the Bonferroni/Dunn procedure when the F ratio for the ANOVA was significant at $p < 0.05$. In the following table, IS represents "Indirect Sound."

The main effect of the factor "Indirect Sound" was significant (p<0.01), as was the the main effect of the factor "Visually Impairment" (p=0.046). These results suggest that blind people perceive distance to be significantly longer than so sighted people. Furthermore, the Bonferroni/Dunn test revealed the trend that there were significant differences among different indirect sound conditions around -40dB but no significant difference near the complete reflection condition (-0dB) nor near the condition of no indirect sound condition. There was no significant interaction effect.

These results indicate that blind subjects estimated or tended to estimate the time-to-contact significantly longer than did the sighted subjects. However, there was no significant difference between the blind and sighted subjects with respect to sensitivity toward each experimental factor, the speed of a moving sound source and the gain level of indirect sounds.

Therefore, in future experiments, we must pay attention not to the change in speed and gain level but to how blind people use direct and indirect sound under a certain speed and a certain gain-level condition. These results also agree with the fact presented in our previous study that locating errors become larger when a sound source moves more slowly [5,6]. Furthermore, as there was a significant difference among different indirect sound conditions around -40 dB, it appears that experiments can be performed under these conditions in order to examine what clues can be used to perceive crossability.

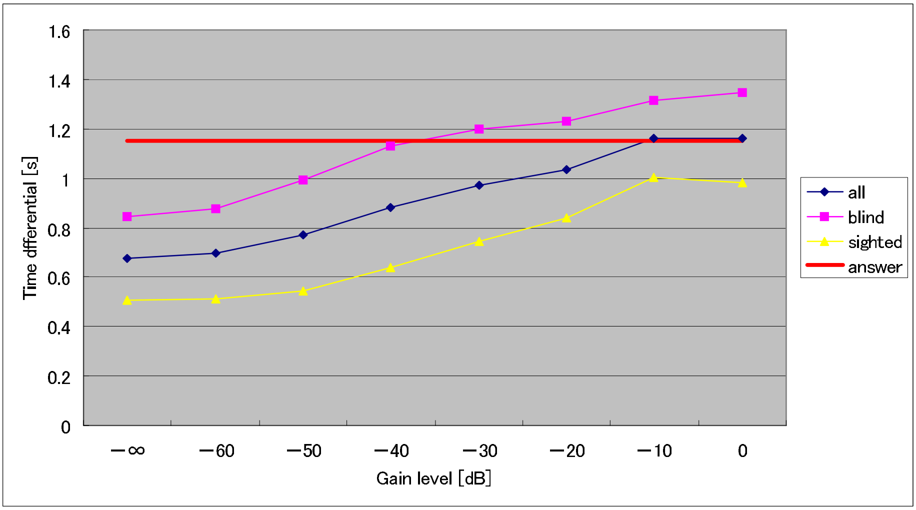


Fig. 4. Accuracy of locating moving sound source vs. change of indirect sound conditions

Table 2. Two-way ANOVA table for changing gain levels of indirect sound

Source	Seq. SS	df	MS	F	P	
	44.06210	24	1.835920	32.31988	0.0000	**
VI	8.12675	1	8.126753	4.42648	0.0460	*
IS	6.93716	7	0.991023	17.44615	0.0000	**
VI*IS	0.171770	7	0.024538	0.431982	0.8812	
Residual	9.543187	168	0.056804			
Total	68.84090	207				

Note: * $p < 0.05$, ** $p < 0.01$

5 Conclusions and Future Works

These results suggest that we should classify all subjects into groups depending on the difference of high or low sensitivity to changes in speed and indirect sound, not on their sight. According to this classification, we should be able to develop training menus that correspond to individual variations of used acoustic clues. We are now developing new training menus following this line of thought.

Acknowledgment

I heartily acknowledge that this research is supported by the Nissan Science Foundation and by the Japan Society for the Promotion of Science Grant-in-Aid for Young Scientists (B): no.16700540 and appreciate their valuable assistance.

References

1. Ashmead, D.H., Davis, D.L., Northington, A.: Contribution of Listeners' Approaching Motion to Auditory Distance Perception, Journal of Experimental Psychology: Human Perception and Performance, **21**(2) (1995) 239–256
2. Brunswik, E.: Perception and the representative design of psychological experiments. Berkeley, CA: University of California Press, (1956)
3. Cooksey, R.W.: JUDGMENT ANALYSIS -Theory, Methods, and Applications-. Academic Press, Inc., (1996)
4. Gibson, J.J.: The ecological approach to visual perception, Hillsdale, NJ: Lawrence Erlbaum Associates, Inc. (1986)
5. Ito, K.: Auditory interceptive timing and familiarity with acoustic environment. Studies in Perception and Action IV, (1997) 83-87
6. Ito, K.: Detection of occluding edges by the use of echolocation. Studies in Perception and Action V, (1999) 52-56
7. Lee, D.N.: A theory of visual control of braking based on information about time-to-collision, Perception, **5** (1976) 437–459
8. Mershon, D.H., Bowers, J.N.: Absolute and relative cues for the auditory perception of egocentric distance, Perception, **8** (1979) 311–322
9. Rosenblum, L.D., Carello, C., Pastore, C.R.: Relative effectiveness of three stimulus variables for locating a moving sound source, Perception, **16** (1987) 175–186
10. Schiff, W., Oldak, R.: Accuracy of judging time to arrival: Effects of modality, trajectory, and gender. Journal of Experimental Psychology: Human Perception and Performance, **16** (1990) 303–316
11. Shaw, B.K., McGowan, R.S., Turvey, M.T.: An Acoustic Variable Specifying Time-to-Contact, ECOLOGICAL PSYCHOLOGY, **3**(3) (1991) 253–261
12. Shiose, T., Sawaragi, T., Nakajima, A., Ishihara, H.: Design of Interactive Skill-Transfer Agent From a Viewpoint of Ecological Psychology, International Journal of Human-Computer Interaction, **17**(1) (2004) 69–86
13. Shiose, T., Ito, K., Mamada, K.: The Development of Virtual 3D Acoustic Environment for Training 'Perception of Crossability', Proccdings of 9th International Conference on Computers Helping People with Special Needs (ICCHP 2004), Paris, France, July 2004, pp.476-483.
14. Shiose, T., Ito, K., Mamada, K.: Virtual 3D Acoustic Training System Focusing on Differences among Individual Clues of Crossing Roads, Proccdings of 2nd International Conference on Enactive Interfaces, Genova, Italy, November 2005.
15. Singleton, J.: Japanese Folkcraft Pottery Apprenticeship: Cultural Patterns of an Educational Institution in Coy ed., (1989)

A Computer-Based Navigation System Tailored to the Needs of Blind People

Bettina Pressl and Manfred Wieser

Institute of Navigation and Satellite Geodesy
Graz University of Technology
pressl@tugraz.at,
wieser@geomatics.tu-graz.ac.at

Abstract. For blind people it is very difficult to orientate in an urban area. Especially in an unknown environment they cannot move independently and need the help of a sighted person. With the development of a navigation system tailored to the special needs of blind people their mobility will increase. Blind people make high demands on all components of the navigation system. These requirements and possible solutions are worked out in a project with the collaboration of blind people. The development aims at a prototype of a navigation system which covers the overall spectrum of navigational components. This comprises an appropriate modeling of the navigational environment, fast routing algorithms generating lists of maneuvers, suitable positioning tools, reliable map matching algorithms for route checking, and finally, adequate guidance instructions.

1 Introduction

The aim of PONTES (positioning and navigation of visually impaired pedestrians in an urban environment) is to develop a demonstrator navigation system for blind and visually impaired people tailored to their special requirements on all system components. The research project at the Institute of Navigation and Satellite Geodesy of Graz University of Technology is funded by the ARTIST programme of the Austrian Federal Ministry of Transportation, Innovation and Technology. ARTIST, the Austrian Radionavigation Technology and Integrated Satnav Services and Products Testbed, is founded to promote the development of applications for the Galileo System which will be Europe's future satellite-based navigation and positioning system. A close cooperation with the Styrian Association of Blind and Visually Impaired People permits the involvement of the user group into the project.

2 The System's Concept

The system's architecture is similar to the concept of an in-vehicle navigation system but is tailored to the special needs of the visually impaired pedestrians.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 1280–1286, 2006.

The components shown in Fig. 1 are a navigable map which includes the entire data for the navigation system and a positioning module to determine the current position of the pedestrian. This absolute position is transformed into a location (i.e., a relative position) by the map-matching module. Finally, the routing and guidance module generates guidance instructions, obstacle warnings, and additional information on the vicinity which are transmitted to the user acoustically via a man-machine interface at the correct time.

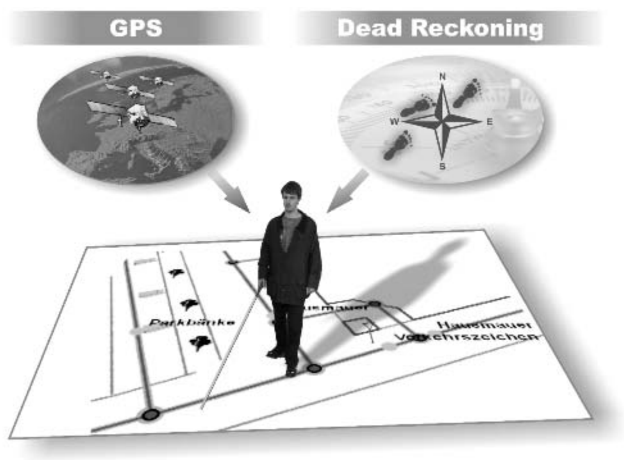


Fig. 1. System Components

2.1 Navigable Map

The database of a navigation system is a digital map that includes the whole geometric, topologic, and thematic information. In case of blind users, it is very important that the map fulfills the following properties: geometrically accurate, topologically consistent, thematically correct, up-to-date, and complete. In addition, the data needs a high resolution to provide blind people with specific information. A lot of landmarks, objects, and attributes for describing the surrounding area have to be included to the detailed path network in the case of blind users to ensure autonomous mobility. Moreover, the map must be able to support routing in the sense of route planning, and guidance. Therefore, the data structure of the vector-type path network consists of nodes (e.g., intersection of pavements) and edges (e.g., promenade, zebra crossing) relevant for route planning and of polygon points (e.g., static obstacles) relevant for guidance. In addition, the digital network includes points of interest like public buildings, shops, restaurants, medical care facilities, etc., and landmarks to provide the blind traveller with important information to navigate without the help of a sighted guide, and get an imagination of the environment close to reality [6]. Fig. 2 shows an example of the path network in the testing area of Graz. This network is modelled with the support of a geographical information system (GIS)

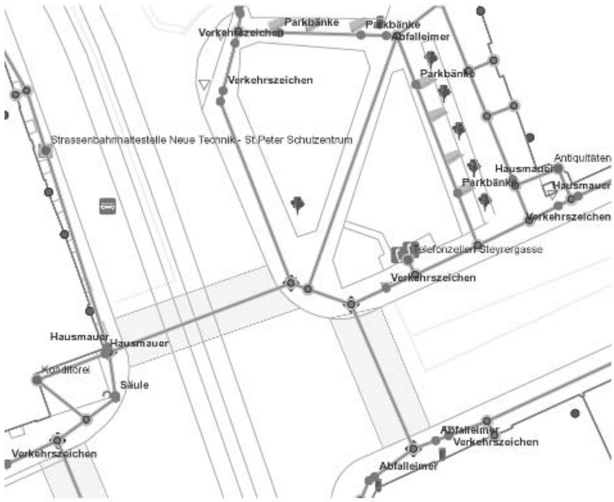


Fig. 2. Path Network

and is stored in a geodatabase. Furthermore, for the route planning algorithm the database structure has to be transformed into an indexed adjacency list to guarantee a high performance [2], Sect. 5.4.2.

2.2 Positioning Module

Compared with conventional pedestrian navigation, the quality requirements on position determination are much higher in the case of blind pedestrians. Accuracy should be at the one-meter level and below. The position must be accurate enough so that objects and locations relevant for guidance instructions are reliably situated within the tactile range of the white cane. Of course, availability must be ensured anytime and anywhere with a top-level integrity. To achieve the above quality requirements, an integrated concept of position determination is indispensable. On the hardware side this is realized by a multisensor system where the determination of the pedestrian's state vector (position and velocity, and to a certain extent even attitude) has to follow the principles of sensor fusion. Hence, appropriate methods of signal processing, i.e., methods of computing and updating (or filtering) the actual state vector are required. In kinematic mode, optimal filtering is achieved via Kalman filtering based on measurements from the multisensor equipment and on dynamic modeling of the motion. Typically, the integrated concept for pedestrian navigation is the combination of satellite-based positioning, e.g., GPS using code pseudoranges, or differential GPS, and dead reckoning (i.e., relative positioning with respect to a known initial position). The latter allows the bridging of GPS data gaps typically occuring in an urban environment [8].

The positioning module is realized using the Pedestrian Navigation Module (PNM, Fig. 3) kindly offered by the Swiss Vectronix AG. This multisensor system

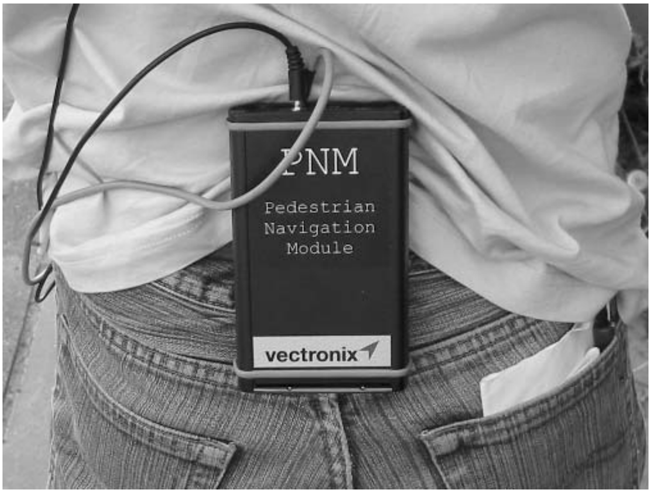


Fig. 3. Pedestrian Navigation Module

includes a GPS receiver for absolute single point positioning, a magnetometer triad and a gyrocompass for course determination, an accelerometer triad for step detection (derived from frequency analysis mainly in the vertical component), and a barometric altimeter for height determination. The associated firmware calculates the DR position and integrates it with the GPS position using a Kalman filter [3].

2.3 Map Matching

Map matching is the process of projecting the determined position to a digital path network held by the navigation system. I.e., the current position is matched to a map point which is directly situated on an edge of the model graph. In other terms, the real position, absolutely defined in a coordinate reference frame, is transformed into a location relative to the nodes of an edge in the graph [5]. This location received in the map is necessary for generating the guidance instructions and obstacle warnings which are communicated to the blind user and depend on topological relations in the map: e.g., an obstacle warning is announced a few meters prior to the occurance of the hindrance, certain maneuver instructions (e.g., “turn left”) are announced immediately before the direction of the path changes, etc. It is worth noting that map-matching calculations are complex and time-consuming, since sophisticated map-matching algorithms perform an edge-to-edge and not a point-to-point matching. Not only the current position is taken into account but also a (partial) “history” of the trajectory. As a consequence map-matching algorithms must also be able to warn of off-route or even off-road situations. In the first case, the pedestrian is still moving on a path within the network but not on the precomputed one, in the second case, he has totally left the path network.

As far as accuracy of the map-matched position is concerned, the results are near to the required values but are still to be improved until the end of this project and in the future. The variance of the position after the map-matching process is around two meters and should be improved towards the tactile range of the white cane.

2.4 Routing and Guidance

Special emphasis of PONTES is placed on the determination of the optimal route from a start site to a given destination. This is realized by a route planning algorithm. One of the most frequently used algorithm is Dijkstra's shortest-path algorithm which is implemented in the system, but, in contrast to a conventional navigation system, the demonstrator does not simply search for the (geometrically) shortest path. Instead, possible safety risks for blind people are evaluated. The various influences are summarized by a specific cost function. This approach allows the routing algorithm to automatically avoid (i.e., circumvent) potential, hazardous situations for blind people. The guidance instructions are based on the route planning module and the route checking task which is permanently monitoring the pedestrian's map-matched position relative to the points where guidance instructions are to be reported. The route planning module generates a maneuver list and tells *how* to guide, route checking decides *when* to guide. Besides, the guidance instructions are divided in four different groups. "Direction instructions" define how to move, e.g., left or right. "Maneuver instructions" tell the user what to do (e.g., turn or go until). Moreover, "objects" are responsible for additional information and give better orientation (e.g., traffic lights). Last but not least, "obstacle warnings" are included (e.g., "attention, there is a traffic sign at the edge of the pavement") [7].

3 Special Requirements

3.1 Detailed Database

Blind people have special needs, because they are more dependent on the navigation system than sighted people. Data of in-vehicle navigation systems cannot be used because of their big scale and less details. It is very important that blind people can rely on the system in every situation. To achieve this, a high accuracy of both the data and the position is necessary, that means an accuracy below the one-meter level. What is even more important, the data have to include lots of details to facilitate an imagination of the surrounding close to reality. In addition, blind people need a lot of objects for orientation especially in unknown areas. During discussions whith blind people, differences occur between old and young people but particularly between people blind by birth and people who became blind later in their life. The latter have more problems to determine their course. Furthermore, it depends on the characters, interests, and fears which objects are needed in the database. Therefore, some persons need more information than other and it is hard to find an optimal solution for all.

To guarantee ideal information outputs for all different demands, the possibility of choices should be made available by the user interface. All objects are divided into object classes and categories, so, each of them could be switched on and off arbitrarily. In other words, obstacle warnings could be turned off, if desired by the user. Moreover, categories of points of interest could be selected specifically or turned off as well (e.g., at everyday routes).

3.2 Guidance and Obstacle Warning

Particular sources of danger are objects which do not touch the ground but reach into the pavement at face level. Such objects cannot be sensed by the white cane. Some examples are postboxes, traffic signs, and gates. A further, precarious circumstance is a bicycle lane at the pavement. This makes it necessary to include (obstacle) warnings besides guidance instructions. For each maneuver, two guidance instructions are communicated to the user, one for preparing some meters before and the other one for executing the maneuver. In contrast to guidance instructions, there is only one warning for each obstacle to avoid too much outputs. The guidance instructions are not dependent on the velocity as in in-vehicle navigation systems since pedestrians move very slowly.

3.3 Safest Route

Concerning route planning, it is not only the shortest route which a blind person needs. Elderly people prefer the safest route. This is realized in the system by a cost function for the route planning algorithm. This function avoids dangerous paths like cebra crossing without traffic lights.

4 Difficulties and Problems

4.1 Communication Interface

A very complex part is the design of the user interface. Blind people have a poor cognitive imagination, their spatial reasoning is sequential. Therefore, it is difficult to transmit the proper information of the sourrounding space to a blind pedestrian. As a result, high requirements on the user interface are needed. There are some difficulties in the realization of an optimal user interface: The first intention of everyone would be the use of earphones [4]. But, blind people do not like to use headphones, because in that case they are not able to hear the background noise of the environment. However, this acoustic information is very important for their orientation (e.g., the traffic noise helps blind people to find the right direction for crossing streets). In addition, speech input is very difficult to be realized and is not yet reliable enough to be used for blind people. Especially, in noisy surroundings problems occur with speech recognition.

An idea for an output device is to use vibrating wristbands. Direction information could be given with the vibration on the left or right arm. Moreover, obstacle warning could be realized by varying the intensity of the vibration with the distance from the object. Another user interface device could be an one-hand keypad with braille output, which supports input and output functionality.

4.2 Temporary Obstacles

A further difficulty is the detection of temporary obstacles like bicycles at the pavement, construction sites, scaffolding, and so on. These objects are a special risk of injury for blind people. To include these objects to the obstacle warnings of the navigation system, object recognition via head mounted camera is needed [1].

5 Outlook

A future project called ODILIA will start in July 2006, when PONTES ends. The aim of ODILIA is solving the remaining problems discussed above, optimizing the user interface, and extending the system. Besides the mobile unit, ODILIA will additionally include a desktop for pre-trip information and training at home. As far as positioning is concerned, indoor applicability of the system and the use of public means of transport like busses or trams will be considered. Another main topic in ODILIA will be optimizing the generation of the path network to expand the system to larger environments. This includes the development of an automation process. Furthermore, the blind person should be able to acquire additional obstacles and points of interest, and include these objects to the database.

References

1. Helal, A., Moore, S.E., Ramachandran, B. (2001): Drishti: An Integrated Navigation System for Visually Impaired and Disabled. IEEE, 2001, 149-156.
2. Hofmann-Wellenhof, B., Legat, K., Wieser, M. (2003): Navigation – principles of positioning and guidance. Springer, Wien.
3. Ladetto, Q., Merminod, B. (2002): In step with INS – navigation for the blind, tracking emergency crews. GPS World, Vol. 13, No. 10, 30-38.
4. Loomis, J.M., Golledge, R.G., Klatzky, R.L. (1998): Navigation System for the Blind: Auditory Display Modes and Guidance. Presence, Vol. 7, No. 2, 193-203.
5. Mayerhofer, B. (2005): Map-Matching in der Fußgängernavigation für blinde Personen. Master's Thesis at Graz University of Technology.
6. Petrie, H., Johnson, V., Strothotte, T., Raab, A., Fritz, S., and Michel, R. (1996): MOBIC: Designing a Travel Aid for Blind and Elderly People. Journal of Navigation, Vol. 49, 45-52.
7. Pressl, B. (2005): Digitale Karte zur Zielführung in einem Navigationssystem für blinde Personen. Master's Thesis at Graz University of Technology.
8. Wieser, M., Legat, K., Hofmann-Wellenhof, B. (2005): Navigation of Visually Impaired Pedestrians - A Tailored Concept for an Urban Environment. Proceedings of the European Navigation Conference GNSS 2005 on CD-Rom, Munich.

A Haptic Interface for an Indoor-Walk-Guide Simulator

Yasuyuki Murai[1], Hisayuki Tatsumi[2]
Nobuyuki Nagai[2], and Masahiro Miyakawa[2]

[1] School of Pharmacy, Nihon Pharmaceutical University
10281 Komuro, Ina-cho, Kita-Adachi-gun, Saitama 362-0806, Japan
murai@nichiyaku.ac.jp
[2] Department of Computer Science, Tsukuba University of Technology (NTUT)
4-12-7 Kasuga, Tsukuba, Ibaraki 305-8521, Japan
{tatsumi, nagai, mamiyaka}@cs.k.tsukuba-tech.ac.jp

Abstract. We are developing a haptic-sensable system to help a blind person understand 3D shapes. As a first attempt we have implemented a pathway simulator which simulates a guiding of a pathway through haptic recognition. If we could indicate a pathway by haptic means to the user, i.e., if we simulate a feeling sensed in his palm and caused by a sliding long cane along the pathway, we believe it might give him an on-site feeling of the pathway. The purpose of this haptic pathway simulator is to help a user with his making a mental map of the pathway. So in the simulator we provide guiding information of the surroundings verbally as well.

1 Introduction

A walk-guide is a serious task for a visually impaired person. As a result of repeated walk-guides he builds a mental map of the pathway. In order to build it he uses all his sensations: his remaining sight, echo and sound landscape, touch of his feet on the ground, smell, breeze and others. The mental map is described best verbally. It contains a detailed description of the clues on the pathway which is usually tedious to the sighted.

In our campus a blind freshman is introduced classrooms, hallways, stairs and major pathways while walking around in the buildings together with an assisting guide helper. As this guidance needs an assisting person, it is hardly possible to repeat the training several times. So students are obliged to remember details of orientation only after few practices.

A walk-guide simulator is helpful in this situation. Such a simulator is reported in [1,2] where a map indicating roads and obstacles (buildings) for walking is presented at the computer terminal and a trainee (user) orients himself in the map by using keyboard inputs under a voice navigation. For example, if he inputs an arrow, then his location in the map is moved towards the direction of the arrow by one unit. However, it is easily understandable that a training using a keyboard might give user a feeling that is far different from one obtained by real walking using a white cane.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 1287 – 1293, 2006.

A few researches are reported in the direction where diagrams and figures are presented to a blind person through a touch-sensable interface in which both-directional communications of input and output between the user and the system are allowed by touch sensable means [3,4].

We report a walk-guide simulator using a touch-sensable (haptic) interface. In our simulator a trainee, holding in his hand a stylus which is captured in a limited-range free three-dimensional space of a haptic device, pushes it against resisting force (computer controls this resisting force) toward the floor for walking simulation. A trainee through the stylus can feel a touch sense similar to one which might be obtained when he is using a white cane while walking on the floor. We describe how our simulator works and report about its availability observed in the experiments.

We borrow material largely from our almost simultaneous report in [5]. However, in this paper we have tried to emphasize the use of haptic devices for the visually impaired.

2 Description of the Simulator

Our hardware consists of a haptic interface and a personal computer (PC). We display the layout of the interior of a room by a simplified 2D arrangement of furniture and pathways in the computer terminal. The user of the simulator recognizes indoor objects (desks, shelves and so on) through a haptic interface feeling as if he is touching these objects through a white cane. He can move around in the room as well in the simulator.

The simulator consists of the three parts:

(1) a Phantom Omni (SensAble Technologies Inc.) which serves as a user interface,
(2) its control software and
(3) the interior layout model (the latter two are implemented in PC).

The interior layout model is implemented in OpenGL in 3D graphics. The user can do a walk-guide training starting from the entrance of the room feeling the response of the haptic interface.

In Fig.1 we show the hardware system organization. An interior layout model is shown in the display of the computer; the display is for a weak sighted trainee or for the developers use: the system can be used without the layout display. The device allows a trainee feel a sense of touching onto the objects created in the computer by a graphic modeler.

To operate Omni one moves a stylus (a pen-shape stick on which two buttons are equipped at the holding position which serve as two buttons of a mouse of computer) whose top is attached to the end of the arm stretching out from the rotator body set on the pedestal.

The 3D movement of the top of the stylus is fed into the input of the 3D model which is built in the computer. By the geometrical relation between the input and the object in the computer model, a haptic response is sent back to the user through the stylus. In this way the user can touch or go around the objects in the model. As the simulator responds with repulsive force against touching, the user can obtain a feeling at the stylus as if he is virtually touching the object.

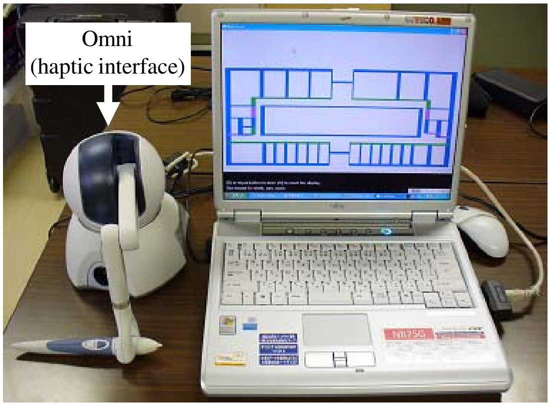


Fig. 1. System organization

2.1 The Walk-Guide Simulator

In the simulator the user can experience as if he is moving around in the layout shown in the display by examining his pathway using a white cane. The walk-guide practice is safe and can be repeated as many times as he likes. The training is expected to be done in a more realistic (simulated) environment than one in which tactile or voice map are used. The layout can be designed according to a real environment or it can be created freely. It is provided as a separate file to the system by a program. The dimension of the whole interior floor of the room is 20 x 20 square units including the four walls (See Fig.2, where (a) and (b) show our room and floor, respectively). An interior object including a wall can be set by clicking the units which the object might occupy (for the programming we used Microsoft Visual Basic 6.0).

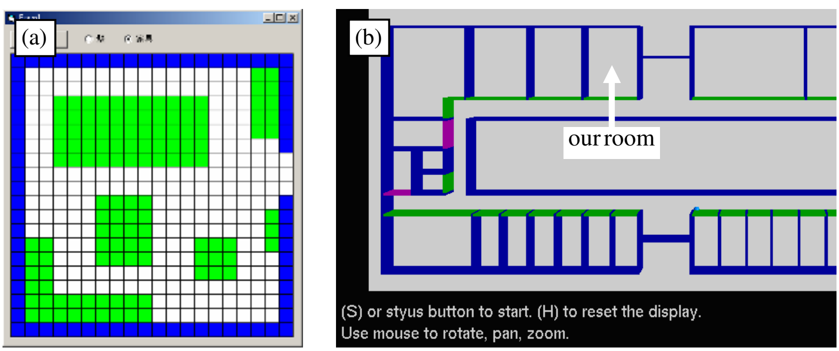


Fig. 2. A room layout

For our interior model we set a size-20-cube as our room, in which we can arrange several objects (furniture) on the floor. Each object is constructed by connecting unit size cubes into an appropriate shape. The room is surrounded by walls with unit size thickness with an entrance which has 3 unit width. We have colored the floor in white, walls in blue and furniture in green. Fig.2 shows a bird's-eye view of our room (from z-coordinate direction).

The layout is displayed on the computer window as a 3D model by OpenGL (See Fig.3). At the same time the layout is fed into Omni through its API (Application Program Interface) and become available in the simulator. The user (trainee), holding Omni's stylus, pushes it against floor (that is, perpendicular to the display surface at his front) and moves on the floor based on the sense of the repulsive force from the floor. If he is confronted with the wall or a desk, a strong repulsive force prohibits him from his further moving. In this way he can train himself a walk around in the room sensing the touch of the floor. Omni allows various settings for the response: stiffness, elasticity, friction and so on. So it can respond pseudo-real responses that are similar to the touch of a real long cane to the user, as well as it can respond artificial responses specially designed to react to, for example, dangerous places; this may be helpful in giving an efficient training.

In Fig.4-a ~ 4-c we show a trainee's current locations. In Fig.4-a he is in the middle of the pathway, in Fig.4-b he is moving along the table, and in Fig.4-c he has made a turn toward us at the sofa. In each figure, a blue cone locates Omni's stylus. The interior layout model constructed by OpenGL can be manipulated directly through Omni's API.

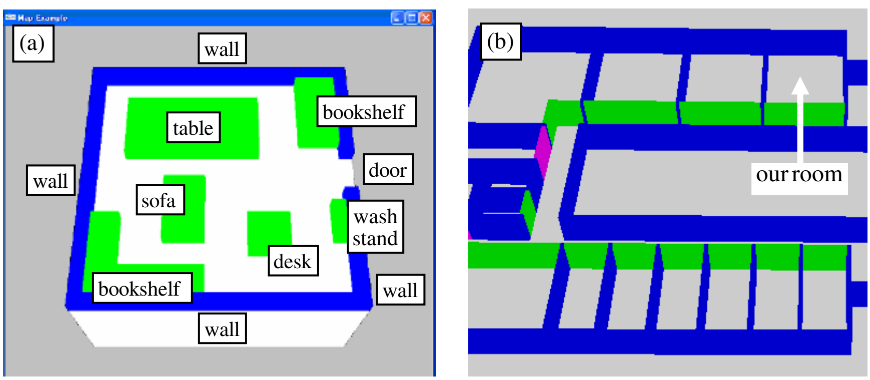


Fig. 3. 3D model by OpenGL

2.2 Program

We have used the C language for programming. The API programs for OpenGL and Omni we have used through SDK (Open Haptics, Academic Edition for Microsoft Windows ver.2.0). We compiled the program by Microsoft Visual Studio.Net 2003. The program size is about 1,000 steps in source codes. Major part of the programming for Omni is to program 3D interior layout model through OpenGL. The OpenGL, originally made by Silicon Graphics Inc., is a 3D graphics library which does not

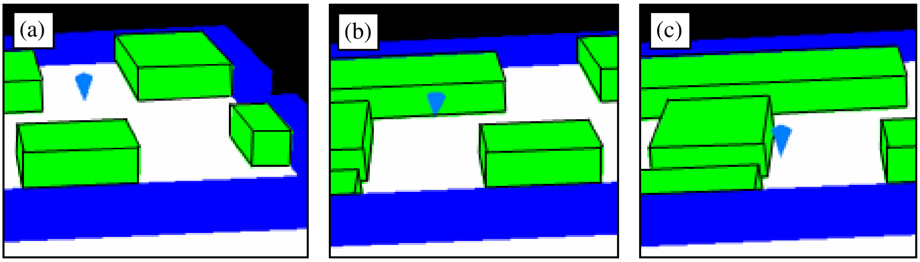


Fig. 4. A walk around in our room

depend on operating systems. Using the GLUT (OpenGL Utility Toolkit) library simplifies 3D programming.

3 Experiments and Discussions

3.1 Experiments

We have tested our simulator by two students (as subjects). The experiment is as follows: we have prepared three interior layout models (See Fig.5-a ~ 5-c). For each model each student spent 10 minutes to walk around in the room. Each student is to recognize floors, walls and furniture. They have given comments on the usability of the simulator. There are several important problems:

1. Every object can be walked thorough easily if the user gives an extra force to the stylus.
2. It is very hard to recognize the real size of the objects.

For the problem 1 it could be adjustable as the current implementation sets zero repulsive force against too strong force in order to protect hardware device. For the problem 2 we need further investigation. In Fig.6 we show a picture of the experiment.

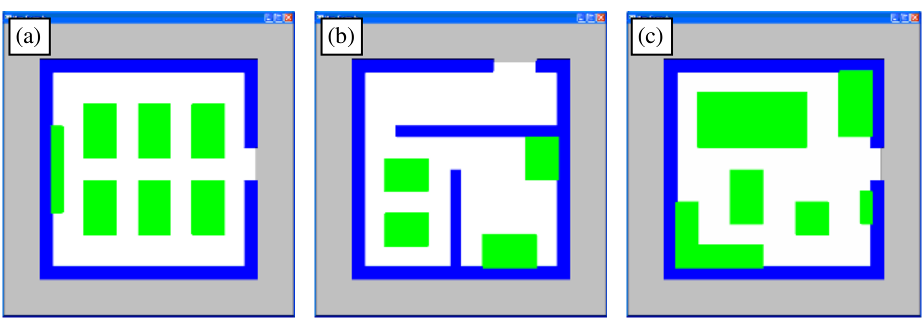


Fig. 5. An experiment of the three layout models

Fig. 6. An experiment scene

3.2 Comparison with a Tactile Map

For a walk guide purpose a user (student) is usually given a tactile map. A sheet of tactile map has several remarkable features compared with our simulator, especially with our layout model.

1. Making a tactile map is simpler than making our layout model.
2. Tactile map is portable, i.e., the user can bring it to the place with him.
3. Tactile map is flexible to use; the user can manipulate all his finger-tips to touch and sense the map. Consequently, he can grasp the global distribution (topology) of the layout more easily. On the contrary, in our simulator the user is allowed to manipulate strictly a point of the stylus.

Nevertheless we see some merits of our simulator.

1. A programmable feature of our simulator makes its potential usage high. For example, we can introduce the user with an automatic guided tour of a route on the layout, i.e., the program automatically guides him a predefined route on the layout with an appropriate voice guidance. Also a guidance by a teacher (a duet walk-guide) is easily realizable; this mode of usage is realized by combining two Omni interfaces connected in master-slave mode through LAN and controlled by a single program; on one Omni the teacher (sighted) guides a route on the layout, the movement of which stylus is transferred to the second Omni, to the student (blind), who follows the teacher holding the second stylus.
2. We can incorporate a rapidly growing computer technology in our system. We have observed that our simulator is in some sense awkward to the user. However, walk-guide simulator with a haptic interface has a rich extensibility due to rapidly growing technology behind it.

4 Conclusions

Our simulator shows that the haptic interface considerably helps a visually impaired person with the training of indoor walking. It is important that he can repeat the training by himself as many times as it is needed. Haptic interface is useful for augmenting tactile or touch sensing which is one of the major devices of sensing for the visually impaired. As our future plan we include:

1. Improving the recognizability of various objects by fitting parameters of Omni: stiffness, elasticity and abrasion.
2. Improving verbal information together with sensable information.

A large size model equipped with many details which reflect real environment would be needed for a fruitful study of our simulator.

Acknowledgment

This research is done under NTUT (National Univ. Corp., Tsukuba Univ. of Tech.) Education Promotion Grant 38: Development of a haptic recognition system for the visually impaired using virtual 3D shapes.

References

1. Ienaga T., Matsumoto, M., Toyoda, N., Gotoh, H., Kimura, Y., Shibata, M.: The production rules and the usability of auditory maps for visually impaired people, Proc. 4th Forum on Information Technology: FIT2005 (in Japanese), No. K-075 (2005) 525-527
2. Komago, K., Koizumi N.: Research on the course guidance using a voice sign digital map, Proc. 4th Forum on Information Technology: FIT2005 (in Japanese), No. K-076 (2005) 529-532
3. Van Scoy, F. L., Kawai, T., Darrah, M., Rash, C.: Haptic display of mathematical functions for teaching mathematics to students with vision disabilities: Design and proof of concept. In: Brewster, S., Murray-Smith, R. (eds.): Haptic Human-Computer Interaction. Lecture Notes in Computer Science, Vol. 2058. Springer-Verlag, (2001) 31-40
4. Tazaki, M., Nagashima, Y., Sakai, T.: An experiment of factorial effect for 3D form recognition on tactile-force feedback display, Proc. 2002 Symposium on Human Interface (in Japanese), No. 2223 (2002) 263-266
5. Murai, Y., Tatsumi, H., Miyakawa, M., Tokumasu, S.: An indoor-walk-guide simulator using a haptic interface, WAC (World Automation Congress) 2006 Proc. 5th Int. Forum on Multimedia and Image Processing, Budapest, Hungary (July 2006), to appear.

Computer Vision-Based Terrain Sensors for Blind Wheelchair Users

James Coughlan[1], Roberto Manduchi[2], and Huiying Shen[1]

[1] Smith-Kettlewell Eye Research Institute, San Francisco, CA USA
{coughlan, hshen}@ski.org
[2] Dept. of Computer Engineering, Univ. of California at Santa Cruz USA
manduchi@soe.ucsc.edu

Abstract. We demonstrate computer vision techniques designed to aid blind or severely visually impaired wheelchair users. These techniques will be used to sense important features in nearby terrain from images collected by cameras mounted rigidly to the wheelchair. They will assist in the detection of hazards such as obstacles and drop-offs ahead of or alongside the chair, as well as detecting veer, finding curb cuts, finding a clear path, and maintaining a straight course. The resulting information is intended ultimately to be integrated with inputs from other sensors and communicated to the traveler using synthesized speech and/or audible tones and tactile cues, supplementing rather than replacing the user's existing cane, guide dog and wayfinding skills.

1 Introduction

Approximately one in ten blind persons uses a wheelchair, and independent travel is currently next to impossible for this population. Conventional blind wayfinding techniques – cane or guide dog – become extremely difficult or impractical in a wheelchair, requiring great physical dexterity and coordination. As a result, independent travel is so difficult that few attempt it, resulting in a widespread lack of awareness of this severely disadvantaged population.

We have begun a research project to develop computer vision technology for sensing important terrain features as an aid to wheelchair navigation. These features include drop-offs, curbs/curb cuts and the shoreline (i.e. edge of the sidewalk bordering grass or other terrain, or adjoining a wall). We are developing computer vision algorithms for interpreting visual scenes to infer this visual information, in real time, obtained from images collected by video cameras mounted to the wheelchair. This information will be communicated to the traveler using synthesized speech, audible tones and/or tactile feedback, and is meant to augment rather than replace the information from existing wayfinding skills. The traveler will use this information in controlling the wheelchair himself/herself (rather than relying on robotic control of the chair).

2 State of the Art and Related Technology

The specific problems of visually impaired wheelchair riders have received little study [3]. Indeed, the only commercial device targeted at this population is a version of the laser cane

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 1294 – 1297, 2006.

by Nurion Inc., mounted on the arm of a wheelchair [2]. The laser's fixed pencil beam drastically limits its "field of view," while four added ultrasonic sensors detect only large, tall obstacles within one foot. Several "smart wheelchair" projects have emerged from mobile robotics research [5], mainly directed at persons with severe physical, cognitive and neurological rather than visual disabilities. However, the robotic approach in general removes control from the user, and operation is restricted to controlled environments [8].

Some technology developed in robotics and autonomous vehicle navigation research may eventually be useful in the design of navigation aids for wheelchairs, but has limitations that prevent it from being adopted in the near future. For instance, 3-D sensing for environmental mapping in robotics is performed using a single or double axis lidar (similar to radar but using laser light rather than radio waves). Although lidars produce very accurate distance measurements, they are still expensive and bulky; moreover, they cannot acquire surface color information, which can be very useful for recognition. Curb detection over short distances for safe driving has been demonstrated at CMU with a laser striper and a calibrated camera, which makes for a very economical system [10]. The problem with a fixed laser striper is that the resulting viewing geometry is very limited, while our task requires the ability to detect features over a rather wide field of view.

3 Research and Methodological Approach

In our experiments, one or more pairs of stereo video cameras are mounted on the wheelchair and connected to a laptop computer to produce range estimates throughout the image. We are currently using the Videre Design MEGA-DCS color stereo system, which produces range maps at several frames per second using a simple but fast correlation-based stereo algorithm.

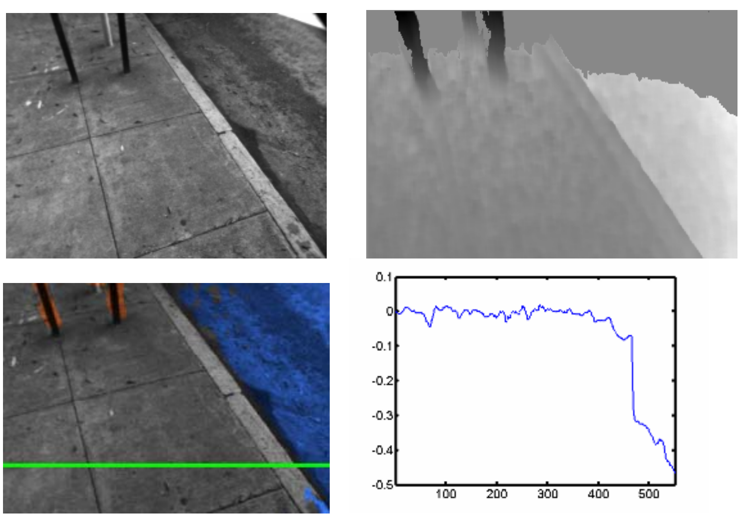


Fig. 1. Stereo input and output. Top, left to right: (a) Image from left camera. (b) Elevation map (darker = higher). Bottom, left to right: (c) Areas classified as above (red), below (blue) and on (no color) ground plane. (d) Graph of elevation across the green slice in panel in (c) is roughly constant except for the discontinuity at the curb.

In order to find important terrain features such as drop-offs and curbs it is useful to first estimate the ground plane in the scene and then use this information to convert the range map into an *elevation map* [7], which indicates the height of points in the scene relative to the ground plane (see Fig. 1b). Since stereo range estimates are often noisy or missing in parts of the image, it is important to develop robust techniques for estimating ground planes and other specific features of interest even when some of the range estimates are unreliable.

4 Findings

In order to improve the accuracy of the elevation map, we have implemented a version of the "V-disparity" algorithm [4] that reliably estimates the ground plane in real time, at a rate of about four frames per second. The algorithm exploits the fact that, in much of the scene, the range is nearly constant along horizontal lines in the image, except at structures that do not lie on the ground plane (see Fig. 1d). This assumption, which our preliminary experiments show to be reasonable in practice, allows the algorithm to reduce the noise in its estimate of the ground plane. We can then use knowledge of the ground plane to detect positive obstacles (visible surface protruding above the ground plane) and negative obstacles, i.e. drop-offs (signaled by discontinuity in the elevation profile, as in [1].

However, the elevation estimates are still sufficiently noisy that additional information should be used to find certain features. For instance, the presence of a curb is signaled by a relatively small elevation discontinuity which may be swamped by noise. We will draw on past work in stereo-based curb and stairway detection [6,9], which uses both elevation and (monocular) intensity discontinuities to identify potential curb candidates.

5 Future Plans

After refining and testing the algorithms for finding the ground plane, drop-offs and curbs, we will design algorithms for finding curb cuts and other ramps and for locating the shoreline. Appropriate modalities for conveying the information extracted by the algorithms – including synthesized speech, audio tones and tactile vibration – will be determined by consulting with blind engineers, who will also help to determine the appropriate algorithm parameters (e.g. how close should a drop-off detected by the algorithm be before its presence is signaled to the wheelchair traveler?). We will also explore possible ways of integrating the algorithm outputs into an overall "clear path indicator," which reports information on where the obstacle-free path ahead is.

Acknowledgments

We would like to thank Bill Gerrey for many helpful discussions and acknowledge support from the National Science Foundation (grant no. IIS0415310).

References

1. Bellutta, P., Manduchi, R., Matthies, L., Owens, K. and Rankin, A.: "Terrain Perception for DEMO III." Intelligent Vehicle Symposium (2000)
2. Gill, J.: Personal electronic mobility devices. Information for Professionals Working with Visually Disabled People. http://www.tiresias.org (2000)
3. Greenbaum, M.G., Fernandes, S. and Wainapel, S.F.: Use of a motorized wheelchair in conjunction with a guide dog for the legally blind and physically disabled. Arch. Phys. Med. Rehabil. (1998) Feb, 79(2), 216-7
4. Labayrade, R., Aubert, D. and Tarel, J.P.: "Real Time Obstacle Detection in Stereo Vision on Non Flat Road Geometry Through 'V-disparity' Representation." Proc. IEEE Intelligent Vehicle Symposium. Versailles, France. June 18-20. (2002)
5. Levine, S.P., Bell, D.A., Jaros, L.A., Simpson, R.C., Koren, Y. and Borenstein, J.: The NavChair Assistive Wheelchair Navigation System. IEEE Trans Rehabilitation Engineering, (1999) Dec; 7(4):443-51
6. Lu, X. and Manduchi, R. "Detection and Localization of Curbs and Stairways Using Stereo Vision", IEEE Intl. Conf. on Robotics and Automation, Barcelona, April (2005)
7. Mandelbaum, R., McDowell, L., Bogoni, L., Reich, B. and Hansen, M. "Real-Time Stereo Processing, Obstacle Detection, and Terrain Estimation from Vehicle-Mounted Stereo Cameras." Proc. IEEE Workshop on Applications of Computer Vision (WACV'98), Princeton, New Jersey, October (1998)
8. Nisbet, P., Craig, J., Odor, P. and Aitken, S. "Smart" wheelchairs for mobility training. Technology & Disability, 5, (1995) 49-62
9. Se, S. and Brady, M. "Road Feature Detection and Estimation", Machine Vision and Applications, (2003), 14:157-65
10. Thorpe, C. et al. "Driving in Traffic: Short-Range Sensing for Urban Collision Avoidance", Proc. SPIE: UGV Tech. IV, Vol. 4715, April (2002)

The Blind Interactive Guide System Using RFID-Based Indoor Positioning System

Jongwhoa Na

Hankuk Aviation University, School of Electronics,
Telecommunications and Computer Engineering, Ko Yang Si, Gyung Gi Do, 412-791, Korea
jwna@hau.ac.kr

Abstract. In this paper, we designed and implemented the Blind Interactive Guide System (BIGS) for the blind person to use in the building. The BIGS uses RFID-based indoor positioning system to acquire the current location information of the user. The system consists of two parts: the smart floor and the portable terminal unit. The smart floor is a floor of a building where each tile of the floor has the passive RFID tag which transmits a unique ID number. The portable terminal unit is an embedded system equipped with an RFID reader as an input device so that the BIGS can get the current location information of the user. Using the preinstalled map of the target floor, the blind person can navigate to the final destination. The prototype is implemented and successfully operated.

1 Introduction

The blind person in most case uses the white cane to navigate the markings for the blinds on the floor inside of a building to find the way. This traditional method is passive in that the blind person must find their way so that if they fail to find the marking they may face some problems. This method may be useful for most blind people. However, when the blind has a problem with the sensory organ or in the case of emergency, the traditional navigation method may not be sufficient for the blind person. In these cases, the active navigation method may be helpful for the blind. The guide dog may be a solution but the dog must be trained at least two years and the expense is unaffordable for many people. The integration of the current technology such as position recognition, embedded system, and the wireless networking technology may provide an artificial blind guide system so that the blind can navigate more easily.

In this paper, we designed and implemented the Blind Interactive Guide System (BIGS) for the blind person to use in the building. The BIGS uses RFID-based indoor positioning system to acquire the current location information of the user. The system consists of two parts: the smart floor and the portable terminal unit. The smart floor is a floor of a building where each tile of the floor has the passive RFID tag which transmits a unique ID number. The portable terminal unit is an embedded system equipped with an RFID reader as an input device so that the BIGS can get the current location information of the user. Using the preinstalled map of the target floor, the blind person can navigate to the final destination. The BIGS is also equipped with the

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 1298 – 1305, 2006.

wireless LAN so that the user may be connected to the server to be monitored by the security personnel of the building. If the user wants the privacy, the user may disable the wireless network so that the server gets no information from the user.

Next, we present the brief survey for the indoor positioning technology in Section 2. The organization of the BIGS is explained in Section 3. The experimental setup of the system and its operation are presented in Section 4. Section 5 concludes this paper.

2 Overview of the Indoor Positioning System

The most demanding technology for the Blind Interactive Guide System is the positioning system. The outdoor positioning can be easily achieved using Global Positioning System (GPS) which became popular in the vehicle navigation. However the indoor positioning is difficult since the GPS cannot be used inside of a building. Currently, there are few candidate technologies for the indoor positioning. First, the WLAN-based positioning uses several Access Points (APs). In this method, the WLAN-enabled terminal unit reads the signal strength between the AP and the terminal to calculate the location of the terminal unit [1]. The WLAN-based positioning system has reported the accuracy of 5m with 90% probability for the moving objects [2]. Because our navigation application requires an accuracy of less than 2m, the WLAN-based positioning system is not appropriate for the navigation purpose.

Another method for indoor positioning is by using the ultrasound transmitters and its receivers to figure out the position of the object [3]. This method has been used for many years in the robot navigation applications and it can provide the best accuracy. However, the ultrasound positioning system cannot be used in public place where people are moving constantly. Since walking people or moving objects may cause an interference with the ultrasound signal, using the ultrasound positioning system is not plausible in our application.

The RFID system consists of tags, a reader with an antenna. The reader retrieves the unique ID information from a tag. Tags are affixed to an object such as goods so that it is possible to locate where the goods are. The reader can read the tag by using two methods; inductive coupling and electromagnetic waves. In the case of inductive coupling, the antenna coil of the reader induces a magnetic field in the antenna coil of the tag. The tag then uses the induced field energy to communicate the data back to the reader. The inductive coupling can have a readout distance of a few tens of centimeters. In the case of electromagnetic waves, the reader radiates the electromagnetic energy so that the nearby tag absorbs the energy to activate the circuit of the tag. After the tag wakes up, the tag reflects back to the reader. Three frequency ranges are generally used for RFID system: low (100~500 KHz), intermediate (10~15 MHz), and high (850~950 MHz and 2.4~5.8 GHz) [4]. The communication range of a RFID system is determined through the output power level of the reader to communicate with the tags and the design of the antenna of the reader and the tags.

In the RFID-based positioning system, the terminal equipment can read the ID of the tags directly to get the position information. The RFID Tags can be taped on the tile of the floor, on the wall or any place desired.

In the RFID-based positioning system, the portable terminal unit is an embedded system equipped with an RFID reader, WLAN, and a speaker and microphone. The reader gets the current location information of the user. Using the preinstalled map of the target floor, the blind person can navigate to the final destination. The WLAN is used in order to connect to the server so that the user may be monitored by the security personnel of the building. If the user wants the privacy, the user may disable the wireless WLAN so that the server gets no information from the user. The microphone is used to input the voice commands such as ‘home’ or ‘supermarket’ and the speaker is used for voice output such as ‘stairs’ or ‘threshold’.

3 The Organization of the Blind Interactive Guiding System

The block diagram of the BIGS is illustrated in the following figure. The system consists of a PDA, the RFID reader, and tags. The PDA is used only for the demonstration purpose since it includes the WLAN, microphone, and speaker. We used the RFID reader/tag. It has the portable form factor, battery operated, and the readout distance of 2m between the tags and the reader. The BIGS maintains a map of the 2-D array of the target floor. The physical ID of the tags allocated each array element of the map so that the physical ID can be translated into the location of the terminal unit. Although the current implementation looks bulky, the final BIGS can be realized using the single board embedded with RFID reader of the size of a wallet.

The major functions of the Blind Interactive Guiding System (BIGS) are as follows: (1) Recognize the current location of the portable terminal unit. (2) Calculate the direction of the terminal unit. (3) Recognize voice input and output for the user interface (4) Monitor the users who are connected to the server for those who needs helps. We explain the details of the BIGS.

Fig. 1. Overall block diagram of the Blind Interactive Guide System. It consists of a HP 5450 iPAQ PDA, M-200 RFID reader from Hitrax, and tags.

3.1 Hardware and Software Organization of the Blind Interactive Guiding System

PDA was iPAQ 5450 from the HP which has 400 MHz Intel PXA250 processor, 64KB RAM and 64 KB ROM, WLAN connection, speaker, and condenser-type microphone [5]. The RFID Reader is MR-100 embedded RFID reader from the Matrix and the tags [6,7]. The MR-100 has dimension of 8.6cm x 5.4cm x 0.8cm and weigh only 70 grams. It uses Frequency Hopping Spread Spectrum using the 902~928 MHz UHF band and read range of about 3m. Note that the range can be reduced by decreasing the power level and using appropriate tags. The RFID tag was the standard carton tag with adhesive from Matrix UHF RFID tags. It has the dimension of 5cm x 10cm and stores 96 bits codes and 16bit CRC codes.

The ID of the tag is sent to the reader using the RF Air Protocol is EPCglobal Class 0, Version 1. Then the data is sent to the PDA via serial communication for further processing. The PDA software performs following functions: It stores the ID and translates into the logical coordinate of the 2-D array of the map. Then it computes the next direction by using the value of the previous location and the current location. The figure 2 illustrates the Smart Floor where each tile of the floor is tagged with a RFID tag which emits a unique ID. This Smart Floor is used in the BIGS as shown in the figure 3. Once the user speaks one's destination, the BIGS calculates the route from the starting location to the destination location and guides the user accordingly.

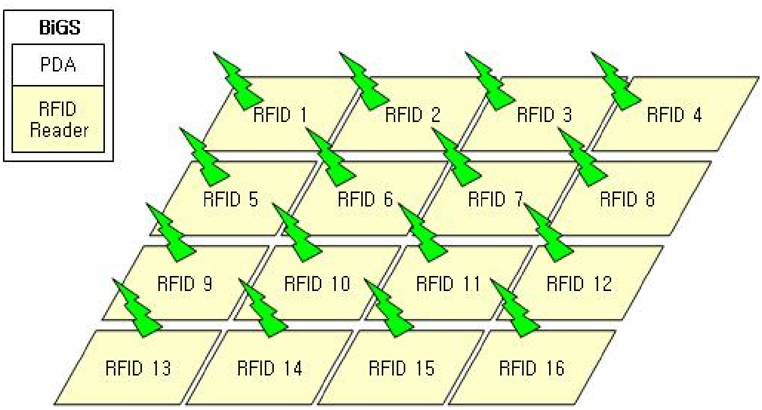


Fig. 2. Block diagram of the Smart Floor for the Blind Interactive Guide System. The space of the Smart Floor is two dimensionally partitioned and tagged with a unique RFID tag.

The flowchart illustrating the detailed operation of the BIGS is illustrated in the figure 4. The first process is getting the voice command from the user stating the destination. The speech recognition processes is executed and compared with the stored speech patterns. Once the destination is identified, the BIGS recognizes its current location and calculates the route. The BIGS outputs the calculated direction using voice until the user arrives at the destination. Once we set the data structure as shown in figure 3, the route calculation can be performed as follows:

```
Let dest_found = false
current location of the user = current.x & current.y
destination location = dest.x =0 and dest.y=0
```

```
While not dest_found
   Decrement or increment current.x & current.y
   If current.x=dest.x and current.y=dest.y
   then dest_found = true
```

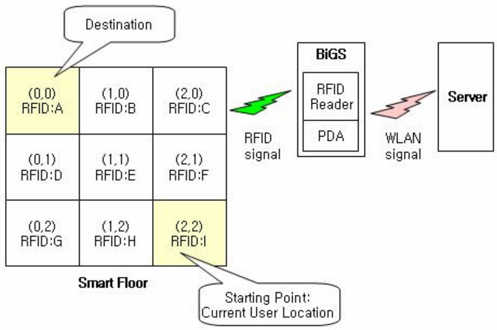


Fig. 3. Illustrated block diagram of the Blind Interactive Guide System

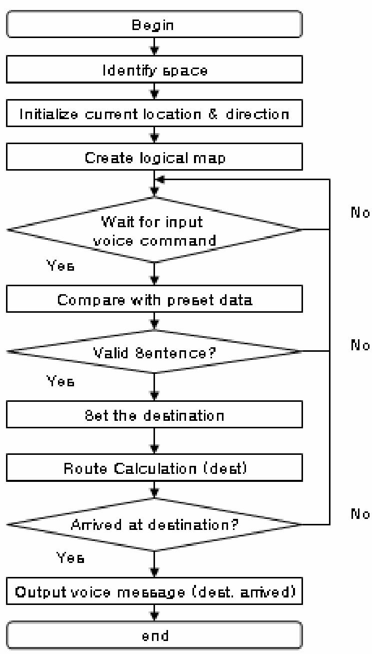


Fig. 4. Main flowchart of the Blind Interactive Guide System is illustrated. After initialization process, the BIGS waits for the voice input from the user. The system directs using voice.

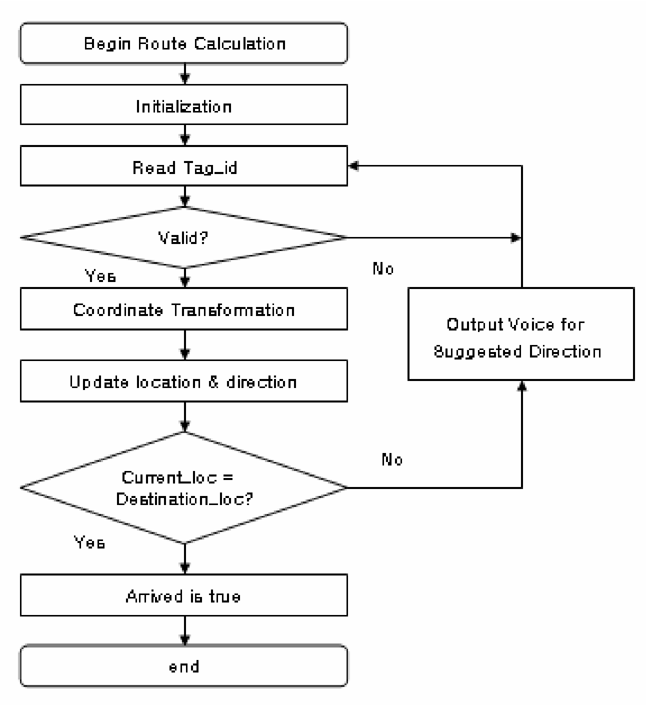


Fig. 5. Route calculation flowchart is illustrated. First, the validity of the tag is checked. Then the next direction is calculated.

4 Demonstration of Blind Interactive Guiding System

The Blind Interactive Guiding System is implemented to demonstrate the functionality of the proposed system. We attached the tags on the long desks to execute the experiments. The distance between the PDA/RFID reader module was around 2m

Fig. 6. The user is starting the navigation. In the right picture, the RFID reader is sitting under the 5450 PDA.

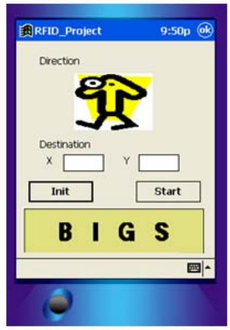


Fig. 7. The user is now on the right direction

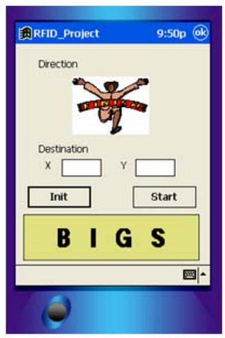


Fig. 8. The user has arrived at the destination. The PDA greets the user with the welcome sign and voice output.

and the picture was taken closely only for the photographic purpose. After voice recognition, the system starts calculation for the route, shows and speaks the suggested direction. The Figure 6 illustrates that the user is recommended to go left direction. Then in the Figure 7, the user goes straight ahead where the destination is located. Finally, in the Figure 8, the user arrives at the destination and hears the arrival message from the PDA.

5 Conclusions and Future Works

We have designed and implemented the Blind Interactive Guide System (BIGS) for the blind. The BIGS is designed by integrating the state of the art IT technologies such as RFID, WLAN, embedded system technologies. The BIGS is implemented using the PDA, embedded RFID reader and tags as a proof of the concept and we have demonstrated its functionality.

We are currently designing a single board version of the BIGS which includes those features described. After the system is stabilized, we will conduct the field test for the usability of the system for the blinds.

References

1. Ekahau, Inc., 12930 Saratoga Avenue, Suite B-9, Saratoga, CA 95070, 1-866-435-2428; see www.ekahau.com
2. Z. Xiang, S. Song, J. Chen, H. Wang, J. Huang, X. Gao, "A wireless LAN-based indoor positioning technology", pp.617-625, IBM J. RES. & DEV. VOL. 48, NO. 5/6 SEPTEMBER/NOVEMBER 2004
3. C. Randell and H. Muller, "Low Cost Indoor Positioning System", International Conference on Ubiquitous Computing, pages 42--48, September 2001.
4. Finkenzeller K, RFID Handbook: Fundamentals and Applications in Contact-less Smart cards and Identification (2002)
5. iPAQ 5400 Users Guide http://h20000.www2.hp.com/bc/docs/support/SupportManual/lpia8003/lpia8003.pdf
6. MR 100 RF Receiver Module Reference Manual http://www.hitrax.co.kr/resource/data/list.htm
7. Matrics Tag Family Datasheet http://www.hitrax.co.kr/resource/data/list.htm

An IT Training Programme for Blind Computer Users – Presentation and Discussion of Didactic and Teletutorial Implications

Mario Batušić[1], Andrea Gaal[2], Joachim Klaus[2], and Mary O'Grady[3]

[1] Institut Integriert Studieren, Johannes Kepler Universität Linz, Austria
[2] Studienzentrum für Sehgeschädigte, Universität Karlsruhe (TH), Germany
[3] Disability Support Service, University College Cork, Ireland

Abstract. The current discussion about the inclusion of people with disabilities into an open labour market is characterised by dominant economic aspects. The unemployment situation of people with disabilities – irrespective of their intellectual level – varies in the different partner countries. The transnational EU LdV Programme IDOL (Inclusion of Disabled in Open Labour Market) aims to open sensibilities and understanding concerning the situation of people with disabilities especially with visual impairment. Its focus is to develop and to implement an IT and psychological training programme for this target group in the different partner countries and in future in all EU countries.

1 Introduction

IDOL (Inclusion of Disabled in Open Labour Market) was a LEONARDO EU programme running from October 2003 to March 2006. Project partners working with people with disabilities in education and training from many parts of Europe (Poland, Bulgaria, Spain, England, Ireland, Austria, Germany) were involved in this project.

As a result of the research of unemployment statistics amongst people with disabilities across Europe, IDOL acknowledged that graduates with disabilities have difficulty making the transition from the educational/training environment to employment. Thus, the IDOL project focused on the needs and requirements of blind people and people with mobility limitations who are about to complete their studies or have already graduated and have to prepare themselves for professional life.

The objectives of project partners and outcomes resulted in:

– Increasing the assertiveness and confidence of people with disabilities planning to enter employment so that they are able to maintain their jobs.
– Increasing the professional qualifications of blind people entering employment through the use of assistive technology and computer skills.
– Making physically disabled people aware of the benefits of having a service dog, especially in the work place.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 1306–1312, 2006.

- Improving the image of people with disabilities in employment so that potential employers and partners in the work place realise the positive input this marginalised group can contribute to the workplace

A long-term result of the project is to increase the number of blind and physically disabled people employed in the open labour market as well as to maintain the jobs of those persons.

A special outcome of IDOL is – to discuss at ICCHP – an IT Training Programme for blind people in how to use assistive technology and develop basic computer skills, that will increase their skills and confidence in entering the open labour market.

2 IDOL IT Training Programme

The training programme is divided into two parts, Level 1 (EATT) and Level 2 (IDOL). Both levels train students in important IT skills with particular reference to skills needed on the open labour market, like using WORD, E-Mail, databases and spreadsheets.

2.1 IT Training Programme Level 1 (EATT)

EATT (Equal Access to Technology Training) is an IT Training Programme for visually impaired computer users that was developed in a former project part funded by the Leonardo da Vinci Programme. The aim of the project is to increase the computer literacy of visually impaired computer users over the age of 35, however the principals of learning new computer skills can be applicable to learners of all ages.

The Introductory IT Course comprises one set of materials for ZoomText users and another for JAWS users. People with a vision impairment in Ireland, the UK, Italy, France and Denmark took part in piloting the first edition of this course.

Correspondently the course is divided into two parts (part 1, part 2) , each consisting of materials for blind users and for low vision students with teachers' and students' materials.

Part 1. This is the motivating part, the "see what you can do" part. The aim is to show the students that they can communicate using a PC with assistive software despite vision impairment. It deals with the following items:

- Starting and stopping the PC
- Windows
- Assistive software
- E-mail
- Word processing
- The Internet

The students are not expected to be able to master specific topics after each session. Success and motivation are keywords in this part of the course.

Teachers' materials
Teachers' materials for blind students using Jaws screen reader
Teachers' materials for low vision students using ZoomText screen magnifier

Students' materials
Students' materials for blind students using Jaws screen reader
Students' materials for low vision students using ZoomText screen magnifier

Part 2. This is a more thorough investigation of the assistive software and the PC in general. The aim of Part 2 is to get the students to a point where they can use their assistive software to work on their own and to qualify them for further IT training. Students are required to master the keyboard before taking Part 2 of the course.

Teachers' materials
Teachers' materials for blind students using Jaws screen reader
Teachers' materials for low vision students using ZoomText screen magnifier

Students' materials
Students' materials for blind students using Jaws screen reader
Students' materials for low vision students using ZoomText screen magnifier

Structure of the Training Programme
IT Training Programme Level 1 is comprised of eight sessions. Each session is planned as follows: 90 minutes work - 15 minutes break - 90 minutes work. The class should meet once a week. It's important that students can revise between sessions either at home or at the training centre. For those students who have access to PCs at home, the software and the PC configuration should be as close to the one used in class as possible. Whilst EATT focuses on developing basic skills for computer users wanting to gain the necessary skills and confidence to begin using a computer competently, the IDOL IT Training Programme introduces more advanced computer skills to the computer user and encourages transition into the open labour market.

2.2 IT Training Programme Level 2 (IDOL)

This Programme is designed to focus primarily on those skills that would be needed in a work environment. Guidelines for students and teachers and in addition Basic Accessibility Settings support the Programme. The training programme is divided into three distinct sections with inbuilt exercises and assignments for each section. There is also a final assignment that participants should complete in order to receive their Certificate of Completion.

Section 1 deals with Reading PDF Documents with a Screen Reader.
Table of Contents
1 What is a PDF Document?
1.1 Why do we use PDF Documents?
1.2 Can a Blind User Read PDF Documents?
2 What You Need to Read PDF Documents?
3 Installing the Accessible Version of Adobe Acrobat Reader
4 Accessing PDF Documents
4.1 Reading and Navigating PDF Documents
4.2 Exercises
4.3 Assignment 1
5 Exporting PDF Documents to Text
5.1 Assignment 2
6 Printing a PDF Document

Section 2 deals with Using Spreadsheets.
Table of Contents
1 What Is a Spreadsheet?
2 Applications for Using Spreadsheets
3 Microsoft Excel Basics
3.1 Starting Microsoft Excel
3.2 Understanding the Layout
3.3 Navigating within a Spreadsheet
3.4 Assignment 1
3.5 Entering Data in Cells
4 Formatting Data
5 Making Selections
6 Assignment 2
7 Using Formulas and Functions
7.1 Formulas
7.2 Functions
8 Sorting a Vertical List of Data
9 Assignment 3
10 Printing and Saving

Section 3 deals with Using Databases.
Table of Contents
1 What is a Database?
2 Starting MS Access
2.1 Choosing What You Want to Do
3 How a Database is Comprised
3.1 How a Table is Made up?
3.2 Creating a Table
3.3 What is a Primary Key?
3.4 Navigating and Switching Views

3.5 Assignment 1
3.6 Assignment 2
4 Forms
4.1 Creating a Form
4.2 The Form Wizard
4.3 Navigating in the Form
4.4 Assignment 3
5 Queries
5.1 What is a Query?
5.2 Creating a Simple Query
5.3 Assignment 4

The IT Training Programme Level 2 is delivered online and includes an online learning environment. A Mailing List allows students to communicate with both teachers and fellow students in a supportive learning environment. A specialized teletutor is to answer more detailed questions and to support the learning process. The teletutor is working at the regional Training Centre and communicates with participants via net.

Structure of the Training Programme
Duration of Level 2 is calculated at approximately 12 hours. The process is dependent on the individual initiative and IT competency of each participant. After finishing the course effectually the trainee will receive the "IDOL Certificate", acknowledging their successful participation in the programme.

2.3 Additional Materials

As an additional help both for trainees and for trainers, IT Training Programme offers a complete list of keyboard shortcuts for the most applications trained through EATT and IDOL training programmes. All keyboard shortcuts for the MS Office applications relate to the Office version 2000. The reason for this choice is that the Screen Reader JAWS continues to have the best support for this Office version. Because there are many users of the Screen Reader Window-Eyes in the partner country Poland, it was decided to include also a list of the most important Window-Eyes shortcut keys.

2.4 Training Centres

All IDOL IT training materials are online and free for use. Disability support centres and job exchange bureaus in Poland, Bulgaria and other partner countries are encouraged to use these deliverables to organize computer courses. Courses will be offered in following training centres:

Austria Johannes Kepler University, Linz
Bulgaria MARIE-CURIE ASSOCIATION, Plovdiv; EUROINFORM, Sofia
England Roehampton University, London

Germany Universität Karlsruhe (TH), Karlsruhe
Ireland University College Cork, Cork
Poland Jagiellonian University, Krakow; University of Silesia, Katowice; Nicolaus Copernicus University, Torun

Discussions among the different partners led to the result, that besides electronic versions of the IT Training, a CD-ROM was produced with programmes in all partner languages. This fits with clients, who have no internet connection.

2.5 Didactic and Teletutorial Network

During the three year programme process it was only possible to develop the different programme materials and to evaluate them, but it was not possible to build up a comprehensive didactic and electronic learning surrounding. The following graphic underlines the interaction between trainer/teletutor, trainees and the internet platform/CD.

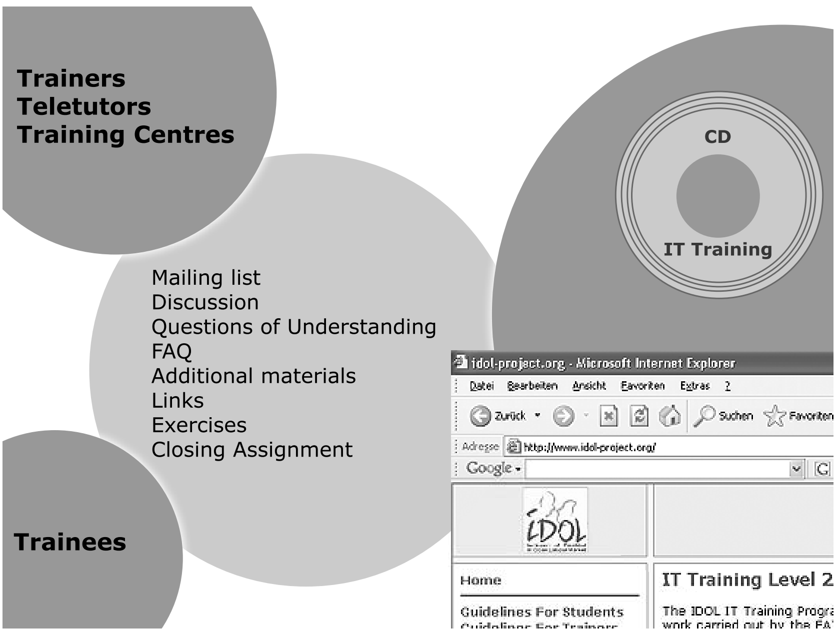


Fig. 1. Network between trainer, trainee and learning platform

2.6 Future Steps

Like all EU Programmes the IDOL project is oriented on sustainability and focuses on building up networks within future training centres. Starting with the centres above all countries are invited to join IDOL and to act with the training centres. A special service will be offered in the near future by developing a training programme for teletutors. Based on the teachers' materials the Study Centre for Open and Distance Learning at the Universitaet Karlsruhe (TH), which has a long experience in media based learning process, will be prepared to include this competence into the realization of the Teletutorial Training Programme. This Programme will be available to all training centres.

3 Conclusion

The current discussion about the inclusion of people with disabilities into an open labour market is characterised by dominant economic aspects. The focus of IDOL on IT and psychological training programmes for this target group is a highly valuable input. Concerning the use of the results and looking for sustainability it will be important to include the training into regional and national programmes. The testing activities evidence the correspondent need.

The unemployment situation of people with disabilities – irrespective of their intellectual level – varies in the different partner countries. The transnational work opened sensibilities and understanding from one to another and was an important input to a common training programme. On the other hand the national partners functioned as transmitter into the national situation. Experts from the partner countries made contact with each other. Therefore the IDOL project can be considered as a contribution to an European partnership and integration.

Usability for All: Towards Improving the E-Learning Experience for Visually Impaired Users

Eva Patrícia Gil Rodríguez, Muriel Garreta Domingo, Jordi Planella Ribera, Magí Almirall Hill, and Llorenç Sabaté Jardí

Universitat Oberta de Catalunya. Av. Canal Olímpic s/n 08860 Castelldefels – Barcelona Spain
{egilrod, murielgd, jplanella, malmirall, lsabate}@uoc.edu
http://www.uoc.edu

Abstract. The new Information Society context entails new risks of social exclusion due to the digital exclusion even when it generates new possibilities for inclusion thanks to accessible Information and Communication Technologies (ICT) and universal design [15]. In this context, we see that e-learning provides a powerful tool towards social inclusion. Yet, we have seen that compliance with accessibility and e-learning standards together with the usage of adequate devices does not guarantee a satisfactory experience for those with a disability [9]. As a result, our project specifically tackles the design of educational technical applications specifically developed for the visually impaired, following a user-centered design perspective.

1 Introduction

New ICT provide us with the key elements to facilitate social inclusion for those who for one reason or another may be experiencing social exclusion. However, if we cannot guarantee universal access and adapted designs to the needs of all different groups, we incur the risk of running into new kinds of social exclusion, known as the digital gap [4]. Of these groups, those with disabilities can easily find themselves excluded from the access to new technologies. This is especially true for people with visual impairment given the great importance images have in this Information Society.

At Open University of Catalonia we have developed several initiatives aimed at facilitating e-inclusion to people with disabilities. We began our work by ensuring that our virtual campus was compliant with accessibility standards and then proceeded to developing output formats for classroom learning materials, specifically designed for students with visual disabilities. For the current project we decided implementing ethnographic methodology to help analyze the overall learning experience, taking into account both the learning contexts and the acquired knowledge.

Through ethnographic methodology we are able to obtain information about the learning and technological requirements for the visually impaired. This information is then used to guide the design of an e-learning course material prototype, allowing us to take into account both usability and psycho-pedagogical aspects. The prototype is then evaluated in a laboratory with both visually impaired users and users without a visual impairment, so that we are not only validating the prototype for the visually impaired, but also to help us understand to what extent a universal design in possible, a design that is good for both groups of users.

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 1313 – 1317, 2006.

2 New Perspectives About Disabilities and the Use of ICT

Traditionally, studies have regarded disabilities from a medical perspective, from which disability is understood to be pathological, associated with an individual and, therefore, considers that the problem should be resolved individually. However, various voices have started to opt for a new environmental model [17]. From this perspective, disabilities correspond with determined interactions that lead individuals to certain situations. Therefore certain disabilities can be corrected if we put into play key devices and tools to facilitate interactions like, for example, installing access ramps in buildings, or designing web pages following the standards of accessibility. This environmental model is taking us to a different definition, which considers disability a conjunction of restrictive conditions that emerge because society is not capable of satisfying the needs of those affected by a disability [13].

New concepts of disabilities enhance the importance of the psychosocial processes and the contexts associated with the problem of the disability. To better analyze the context of the disabilities we need to use the right tools. For example, ethnomethodologies allow considering psychosocial issues, the context and the experience of the user. Ethnomethodology has already been used in a design perspective focused on the user [3], but to date, it has barely been applied to the area of accessibility. This psychosocial dimension has also been backed by what is known as the paradigm of the social perspective of disabilities [1,2,13]. From this perspective, new technologies play a relevant role, as they permit to dilute disability while allowing this group of people to maintain a high level of quality of life.

From this point of view, we support that e-learning and training by means of new technologies can be a key instrument to bridge the digital gap, providing that the instruments are designed adequately so as to facilitate the learning of people with disabilities as a collective. Furthermore, promoting knowledge of the use of these technologies will considerably improve digital inclusion.

3 The Importance of the User Experience

Nowadays, when people with any type of disability navigate the internet they do not have the same advantages as those without a disability. We believe that this is one of the reasons why this group of users has had numerous difficulties accessing long distance learning programs. For several years now, we have addressed this issue within our University trying to identify solutions to these difficulties. Following we describe briefly the changes that we have carried out with this goal in mind, and which has lead us to the project here presented.

Our work began with an accessibility analysis in order to ensure that the virtual campus complied with accessibility standards. There are several actions on the Internet aiming to standardize accessibility. From these, two key norms have to be considered, the “section 508” [16] and the Web Accessibility Initiative (WAI) [17]. This last norm is divided in three levels: A, AA, AAA. The level AA of WAI is similar to “section 508” as shown by the Thatcher analysis [18].

The AA level project took the following steps to accomplish its mission: first, we taught a course on accessible design to in-house web designers; second, we chose a set of pilots to apply the W3C recommendations for the design of the web pages; third, we evaluated the accessibility of the obtained pages. Finally, we progressively made accessible all remaining web sections and we created a style guide for new developments.

Following the accessibility project, we started working on our classroom learning materials by identifying and defining different output formats, one of them which had to specifically be designed for visually impaired students. The Digital Audio-Based Information System (DAISY) [5] is a new technology that helps develop and distribute books and contents. With DAISY it is possible to use the Digital Talking Books (TDB) in order to meet the needs of visually impaired people. Books in DAISY format grant greater speed of reading and greater easiness of access to different sections of the book. This format can be read by small portable devices as well as by a personal computer with special software, such as the system that plays digital talking books on a PC, analyzed by Morley [10].

We then followed a user-centered design process for creating the different output formats for learning materials. We began by gathering quantitative and qualitative information to help us build user profiles. The resulting profiles were the basis for the recruitment screening for user testing. We ran ten user tests per profile identified at which we used a high-fidelity prototype of the materials where students would reproduce the way they usually work with these interactive materials. The user tests were the basis for the interface improvements, based on which we built a new prototype to run a second set of user tests. The results of the second iteration showed a design improvement of more than 50%, compared with the first set of tests.

Both projects were extremely successful and helped us realize that the use of standards and devices does not guarantee by itself a satisfactory experience for the user with a visual impairment or disability, nor does it guarantee that the applications are going to have optimum conditions to reach their formative objectives. For example; a person that is visually impaired can use a screen reader to access the internet; a device that can only give access to those pages that have been designed following the standards of accessibility. However, those standards do not guarantee that the navigating experience through certain pages is going to be satisfactory [9], mainly due to the fact that screen readers read in a linear format the contents of a webpage. Even if this linear format is optimum for an in depth lecture of written material, it is not adequate for hypertextual material. Clearly the wording of web pages for sighted people requires different strategies, more holistic and global than the ones used for written materials.

On the other hand, the processing of the information depends on the way the information is accessed. For example, Braille gives a far more effective comprehension of a text than audio for the visually impaired [8] and this ought to be taken into account in the design of virtual learning materials and e-learning systems. As a result, to enable access to virtual platforms of learning for the visually impaired, it is necessary to recognize the contexts in which these type of students carry out their learning and teaching, as well as the cognitive process that intervenes in their strategies at the time of learning the subject. It is also essential to study the appropriation that they have of technology [10]. This allows us to better their experiences as users, and of course improve the process that they follow to carry on with their plans of learning and the formative objectives that they develop.

4 Usability for All

The general aim of our project is therefore twofold; on one hand we aim to use ethnomethodological skills to obtain basic knowledge about the use of ICT and the requirements of visually impaired users, on the other hand, we analyze the processes and contexts in different learning scenarios. As demonstrated in previous studies.

[7,11], such information can be translated into requirements of the functionalities of devices, and strong recommendations for the design of multimodal interfaces that take into account both the usability and the pedagogical objectives of the educative contents.

Therefore, the project is divided in two main phases: the theoretical and information gathering phase and the prototype development and evaluation phase; the former one guiding the design of the prototype. The following tasks form the initial phase:

- Creation of a theoretical framework regarding the impact of new technologies in the education of visually impaired people.
- Analysis of sociodemographical characteristics of visually impaired people in Spain.
- Quantitative study of the level of education and learning needs of visually impaired people in Spain, their usage of ICT and other relevant variables identified by the theoretical framework.
- Analysis of the technological requirements and user profiles of visually impaired people in the different learning scenarios.

During the second phase, which aims to apply the results of the theoretical framework and guarantee the usability and the pedagogical objectives of the learning material, we will design a prototype of a learning material designed specifically for this collective. This prototype will be assessed in a usability and accessibility laboratory specific to evaluate technological applications destined for visually impaired people.

The portable usability laboratory consists of a PC with TechSmith's usability software and a video camera. Morae Software records the user speech and face, along with the computer screen and the user interactions with the tested application. Observers can follow the evaluation in real-time from other computers. In order to make the laboratory accessible, we will add a screen reader, a screen magnifier, a Braille printer and a reader of digital talking books.

We carry out a two phase in-context evaluation, with prototypes of low and high fidelity that take into account visually impaired users as well as sighted users. By doing this, we can study whether a universal design is possible or if the applications have to be designed specifically for the characteristics of the specific groups of users. As an overall outcome, this project will allow us to obtain a theoretical basis to advance in e-inclusion and digital literacy without barriers.

References

1. All (2003). DisCapacitados. La reinvención de la igualdad en la diferencia. Barcelona: Bellaterra
2. Barton, L. (1999). Difference and difficulty: insights, issues and dilemmas. Sheffield: University of Sheffield. Department of Educational Studies

3. Beyer, H. y Holtzblatt, K. (1998). Contextual Design. San Francisco: Morgan Kauffman Publishers
4. Castells, Manuel (2003) The Internet Galaxy : Reflections on the Internet, Business, and Society. USA: Oxford University Press
5. Digital Audio-based Information System , `http://www.daisy.org`
6. Francis Theofanos, M. & Redish, J. G. (2003). Guidelines for Accessible—and Usable—Web Sites: Observing Users Who Work With Screenreaders. Interactions, nov. des. 2003, pp. 37 – 51
7. Gil, E., Navarro, R. y Casanovas, J. (2005) Una ilustración del uso de observación participante en el diseño centrado en el usuario: cómo estudiar el diseño de aplicaciones intuitivas de banca móvil específicas para el colectivo de inmigrantes. Interaccion 2005. Granada: Spain
8. González, L. (2004). Acceso a la información y comprensión de textos: Comparación de rendimientos entre personas ciegas y videntes. II Congreso Virtual INTEREDVISUAL sobre instrumentos de acceso a la comunicación, la educación y la cultura de personas ciegas. Santiago de Compostela (España)
9. Lawton, S. & Grossnickle, M. (2005). Accessibility in the User-Centered Design Process. Available on `http://www.uiaccess.com/accessucd/overview.html`
10. Morley, Sarah. Digital talking books on a PC: a usability evaluation of the prototype DAISY playback software. Proc. of the third international ACM conference on Assistive technologies 1998, p.157-164
11. Myers, M. (1999) Investigating information systems with ethnographic research. Communications of the association of information systems, Volume 2, nº 23
12. ONCE (2005). Pautas para el diseño de entornos educativos accesibles para personas con discapacidad visual. ONCE. Dirección de Educación. Grupo de Accesibilidad. Plataformas Educativas
13. Planella, J. (2003). Subjectivitat, dissidència i discapacitat. Pràctiques d'acompanyament Social. Barcelona : Claret.Slatin
14. Salmerón, L., Cañas, J.J., Kintsch, W. & Fajardo, I. (2003). The Effect of Navigation Strategies on Hypertext Comprehension Research. Manuscript submitted for publication
15. Shneiderman, B. (2000). What is Universal Usability? Communications of the ACM. Vol. 43, n. 5, pp. 84 – 91 Section 508, `http://www.section508.gov`
16. Slatin, J. (2005). Boundary Objects: The Role of the Web Content Accessibility Guidelines 2.0. Fundamentos Web. Usabilidad, Accesibilidad y Estándares. Gijón (España)
17. Thatcher, Jim. Compare WAI-section 508, `http://jimthatcher.com/sidebyside.htm (2003)`
18. Web Accessibility Initiative (WAI). Available at `http://www.w3.org/WAI`

Issues in Implementing Awareness in Collaborative Software for Blind People

Jaime Sánchez and Nelson Baloian

Department of Computer Science, University of Chile
{jsanchez, nbaloian}@dcc.uchile.cl

Abstract. There is no doubt among the members of the CSCW community that awareness is a key issue in the design of successful collaborative software. In many systems awareness mechanisms have been implemented through displaying graphic information over the system's interface. However, this strategy does not apply when the end users of the system are blind people. In this work we report the problems we encountered when implementing a collaborative game for supporting the learning of music and sound by blind people when trying to develop effective awareness mechanisms. The preliminary results have helped us to be "aware" about some characteristics awareness mechanisms should have for blind people which are not as prominent and problematic for sighted people.

1 Introduction

Many authors recognize that awareness is a cornerstone for designing effective collaborative software. One of the main goals of awareness is to give relevant information to the participants for the completion of a shared task (Gutwin and Greenberg, 1996). This has been recognized as a key issue to help participants to shift from working alone to working together (Collazos et al., 2002). Borges & Pino (1999) also mention that awareness mechanisms are crucial for group interactions. In another work, Gutwin and Greenberg (1996) state the importance of the awareness mechanisms to give up-to-the–minute knowledge of the situation, since most of the collaborative systems implement dynamic changing workspaces. In most cases, the awareness information is given graphically. However, graphical information is obviously not suitable for implementing awareness in groupware for blind people. Although there have been many efforts made for implementing Computer Supported Learning software for blind people, there is almost no information about the development of collaborative software for them. An exception is the work reported in (Sánchez et al., 2004) which presents a game involving sighted and non-sighted people. However, we think that people with visual disabilities can benefit as well as sighted people of this kind of software, not only for learning a certain subject matter such as mathematics, geography, and physics but also for developing their collaborative skills. In fact, traditional education for people with visual disabilities is aimed at developing individual skills only.

Software for people with visual disabilities tends to transform graphic information to a haptic format or audio. AudioDoom (Lumbreras & Sánchez, 1999) allows blind children to explore and interact with virtual worlds by using spatial sound. The game was based on the traditional Doom game where the player moves through corridors

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 1318 – 1325, 2006.

discovering the environment and solving problems simulated with objects and entities that inhabit a virtual world. This proved that sound-based virtual environments can help to develop tempo-spatial cognitive structures of blind children. Roth et al., 2000 presents a game for audio concentration. It consists of pairing different levels of geometric figures, basic, and derives. They constructed a bi-dimensional sound space to represent geometric figures graphically. This concept allows graphic representation such as icons to be represented by the perception of moving sounds in the spatial plane. Morley et al. (1998) through the use of non-speech sounds in a hypermedia interface observed that blind users developed special way of navigating through a known environment and represent spatial structure with cognitive difficulty. VirtualAurea (Sánchez, 2002) is a spatial sound tool editor that can be used by parents and teachers to design a wide variety of spatial maps such as the inner structure of a school, classrooms, corridors, and diverse structures of a house. Users can integrate different sounds by associating them to objects and entities in a story. Ressler & Antonishek (2001) designed integrated active tangible devices such as forcefeedback with a synthetic environment to support collaborative interaction between users.

Since there have been very few developments in collaborative software for blind people we refer to awareness issues developed by authors of software for sighted people. However, we think that if the way of implementing awareness for sighted people may be not feasible for blind people, the principles remain almost the same especially those referring to the knowledge users need to be able to collaborate in a better way. Awareness has been defined as “an understanding of activities of others, which provides context for your own activity” (Dourish & Belloti, 1992). It is especially important to remember that “awareness is a state of mind of a user…while awareness mechanisms are techniques employed by a system to achieve this state of mind” (Sohlenkamp, 1999).

Examples about how graphical information has been used for implementing awareness in CSCW system can be seen at (Gross & Prinz, 2003) and (Moran and Anderson, 1990). However, sound has already been used successfully for implementing awareness in systems for sighted people as reported in (Isaac et al., 2002) and (Cohen, 1994).

In this paper we describe software that implements a collaborative game for people with visual disabilities. We focus the description of this software on problems we found while trying to implement awareness in collaborative software for this type of users. We present our preliminary work on the awareness issue for collaborative software for blind children by highlighting the “awareness” issue in the community.

2 Camino Musical (Musical Path)

In the design of the game participated experts in education of people with visual disabilities. They elaborated the metaphor of the game in order to motivate the students. The metaphor is inspired in the life of Beethoven. The scenario of the game is the year 1826 when the musician suffers a pneumonia which triggers health difficulties. The players must help him to keep his music alive and learn about the musician. They have to try to reproduce his music with different instruments by recognizing different basic parameters of the sound: 1) tone, which corresponds to the note of the sound played by an instrument, 2) duration of the sound 3), volume (high, medium, low), and 4) instrument. Thus, the software teaches the user notions of rhythm and musical scale.

The game starts with the participants choosing their instruments. There are a number of instruments available and everyone has to choose a different one. If an instrument is chosen by one player it cannot be chosen by the others. The rest of the game is divided in two stages in which the players have to recognize the properties of a sound played by the instrument they have chosen. As they progress recognizing the properties of the sounds listened (like, tone, volume, and duration) they add more notes to the melody supposed to be reconstructed in order to play it with the rest of the group. They will have to recognize only one property of the sound in the beginning and more than one at advanced stages.

As soon as the instrument responds to questions learners can hear its improvements and suggest changes through face-to-face dialogs in the case they are playing in the same room. A virtual chat based on voice communication is still not implemented at this stage of development.

The first stage considers three types of questions: to recognize the intensity of sound, to recognize the pitch of the sound, and to recognize both parameters simultaneously. Sounds in this stage correspond to only one musical note. The second stage is conformed of recognizing rhythmic series. The goal in this stage consisted of learners to identify time (semibreve, half note, quarter note, and quaver) within a musical fragment. While learners recognize figures they have to place them in the corresponding order within the composition in such a way that they play in the same order as in the example provided.

At this stage the control panel is conformed of four time figures. They are placed on the inner border of the screen. The associated sounds to each figure are the interpretation of the figures performed by the selected instrument (by the learner) from the beginning of the game.

3 Awareness Issues in Collaborative Software for Blind People

The provision of shared virtual spaces is considered a facilitator of diverse processes between people that work in groups because it supports externalization that plays an important role in the organization and knowledge creation. This implies to support the transition from the tacit and individual to the group. As we stated above awareness mechanisms mostly implemented through visual elements are not effective in cases were users are blind. In such a case these interface awareness elements should be implemented through sounds.

Our application considers that relevant information for a user concerning the other users refers to what instrument others have chosen, in what stage of the software are them, when a player chooses an instrument, and when the player is focused on to perceive. A possibility is to provide sound-based information each time a change happens, that is, when a player choose an instrument and when passes through another stage. However audio information received when state changes are happening may have at least three drawbacks: First, the opportunity of the information. Audio information can arrive when the player is centered on to perceive other sounds then the information can be unnoticed or may alter the work on sound recognition. Second, the retention of information. Due to the fact that this information is volatile it could be not correctly retained by the player. Then it could probably happen that even though the information is perceived opportunistically by users they may not be able to remind it when taken decisions. Third, the validity of the information. Even though the information arrives opportunely and can be retained its validity is loosing throughout the time because the state of the game is changing. This should be reflected on new messages. This can impact drastically the sensation of awareness that the user can have. If the user perceives that understand what is happening to the other players, what they have done, in what stage of the software are them, etc.

Due to the aforementioned reasons the awareness information in this case should be given through audio but only when the user requires it to guarantee the opportunity. This information should reflect accurately what the user needs to know and should be short to promote its retention. To do this it is better to have audio information (earcon) clear and concise instead of an explanation through voice. Finally, the user should ask for the information anytime and it should reflect the up-to-date state.

In our system the awareness information was implemented in the following way: The user can require information about what instruments are in the game (which one were selected). To do this he can press a specific keystroke (F1) to receive information about each instrument at the time each time a keystroke is pressed. The system provides the information in the same order the instruments were selected and in a circular way. The information is a small music piece played by the instrument and its name at the end. We do not believe that it is important to know who is behind each instrument. To know in what stage is a player of an instrument, first, the player should select the instrument following the procedure mentioned and then press F2. After this the player receives a piece of music that has been constructed by other player of the instrument and the piece of sound that is actually analyzing.

4 Testing the Software

4.1 Usability Evaluation

We implemented diverse usability testing to the collaborative software. Two different evaluation methods were applied: usability of interface elements and user acceptance. The evaluation was applied five times during six months to 5 legally blind (three with residual vision and two totally blind) learners ages 10 to16, one girl and four boys, from 5th to 8th grade in the School for Blind Learners Santa Lucia, Santiago-Chile.

The first usability testing was applied to the multimedia resources of the game to select sounds and images to be implemented in the software. Learners identified what resources functioned better to and helped us to design an interface with icons that: convey meaning, be clear to them, avoid ambiguities, and be attractive. Most instruments and figures were clearly identified by learners. However instruments with many details confounded the learners. For instance, the contrabass in some cases was recognized as a violin or a guitar. We also had to improve the size of images in the interface for learners with residual vision. Bigger size icons and images worked very well for them. Colors were also tested and the best results were the combination of red over yellow and blue over yellow.

The second usability testing consisted of improving the interfaces of the game and polishing some deficiencies for the users. We applied usability questionnaires with questions such as do you like the game?, what instrument is playing on the screen?, what colors are you perceiving?, what keyboard strokes do you use and what are their purposes?, what do you would change to the game?, do you want to add something more to the game?. The third usability evaluation consisted of testing the screen with the selection of properties of the sound. For this we used an acceptance questionnaire for end-users.

The fourth usability testing considered testing the first stage of the game. This stage consisted of identifying intensities that can be high, medium, and low. They are randomly generated by the software. Each time the instrument plays, the learner had to identify the intensity and a feedback is provided. After this a new sound is produced by the instrument. For this testing we used an acceptance questionnaire for end-users. The fifth usability testing consisted of evaluating the interface of the menu for selecting instruments and the interfaces for the first and second stage of the game. The menu for selecting instruments contains two navigation buttons that allow moving back and forward through the list of instruments. To move through the menu learners use left and right arrows. To select the instruments they use the space bar. When moving through the menu the next or previous instrument can be shown. A melodic line interpreted by the shown instrument informs to the blind player about the type of instrument selected. When doing this a new interpretation of the instrument is played and a voice says "you have just selected an instrument". In the other clients a voice informs about the selected instrument. If there are players that have not selected instruments the game "waits" until all of them choose an instrument. Meanwhile each time players having a selected instrument press a keystroke they are informed about the players left to start the game.

4.2 Results

Interaction. We evaluated the importance given by learners to learning with the software. We considered the following statements: "I will recommend the software to other players", "I have learned with the software", and "The software allows me to know new things". Blind learners assigned higher scores to the statements showing that they value highly the learning of music contents through the interaction with the software. The statements related to the use of the software were: "I felt controlling the software", "The software is easy to use" and "The software adapts to my pace". At the beginning learners assigned lower scores to this parameter evidencing the complexities of user to manipulate and interact with the software. However once they understood the software the quality of the interaction increased.

Collaborative skills. The last testing included a questionnaire to evaluate the collaborative behaviors of learners and how the software helps to reinforce these behaviors. The following figures consider two basic aspects of collaboration. Figure 1displays positive interdependence through six statements with scores from 0 to 10 points: "I know clearly the objectives", " I know the importance of my role to attain the final goal", "I need to interact with others to attain the final goal", "The fact that each one has to develop specific tasks allow me to value the capacities of my peers", "I feel that the achievement of one of my peers is the achievement of all of us", and "The audio stimuli given by the software when solved correctly a task makes me feel happy even though I was not the one that implemented correctly the action". The average score of learners with residual vision was lower than the average of blind learners especially in terms of positive interdependence. Blind learners assigned more value to positive interdependence statements after collaborating with the computer-based game.

Figure 2 displays the average of the assigned score to personal and group skills answering to the following statements: "I respect the other's turn", "I help to my peer only when a help is required", "I give the necessary time to my peer to answer my questions". "I don't like that my peer does not answer to what I ask or say", "I evaluate my answers and the one of my peers to learn and understand what is going on in the program". Learners with residual vision assigned lower scores than blind learners to these statements. Blind learners really value personal and group skills as a result of interacting with a collaborative computer-based game.

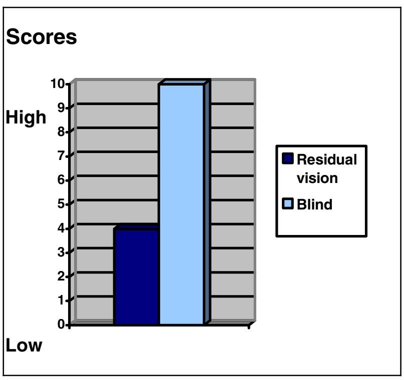


Fig. 1. Positive interdependence

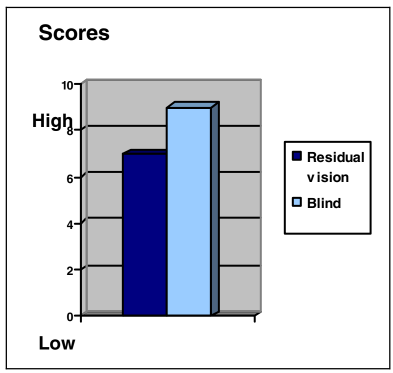


Fig. 2. Collaborative context

Awareness. Two statements during usability testing referred to awareness during interaction with the game: I knew what other players were doing and I knew what instrument had each player. Figure 3 displays the results concerning the first statement. Learners had a hard time figuring out what the other players were doing during the game. Two players could know "sometimes" what other players were doing and three of them definitively knew what other players were doing.

The results of the second statement are displayed in Figure 4. Three learners could "always" know what instrument had each player and two of them mentioned that "sometimes" knew who has what instrument. This is especially interesting because they could know the role and instrument of each participant but could not realize what the other players were doing.

I knew what other players were doing

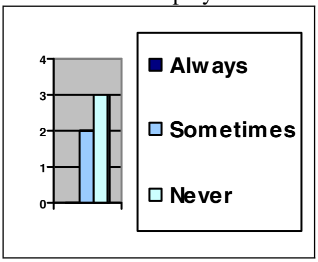


Fig. 3. Awareness of players

I knew what instrument had each player

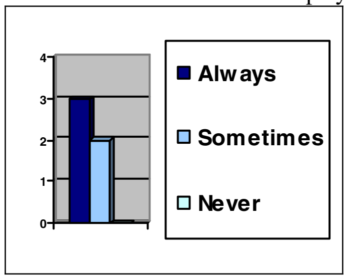


Fig. 4. Awareness of the players' instruments

5 Discussion

This study presents our preliminary work on problems encountered when implementing awareness in collaborative software for blind people. First, we have designed a collaborative application for learning music instruments. Second, we have usability tested the software with five learners evaluating the interfaces for both interaction and collaboration. Third, we have identified some unique problems in implementing awareness in collaborative software for blind learners.

Blind users interacted with the software, recommended improvements in the interfaces and highly accepted the interfaces probably because they mapped well their mental models. During usability we asked them about their knowledge about what the other players were doing but they could not figure out their tasks even though most of them could realize what music instrument had each player. We believe that this behavior can be explained because either the statement was not completely understood or the awareness mechanisms were not correctly implemented in the software.

Our preliminary results confirm our hypothesis about implementing awareness mechanisms in software for blind people. Implementing awareness for this people is not a trivial task. We need more testing to determine which one of the problems stated earlier in this paper, such as opportunity, retention, and validity should be approached differently in order to better fit the requirements of blind people. Our initial data is telling us that principles of identifying and implementing key awareness issues for sighted people should be taken into account but should not be transferred directly when implementing awareness mechanisms for people with visual disabilities. Awareness mechanisms for these people imply more complex problems that must be considered when designing collaborative software for them.

The next task is to fully understand the awareness needs of blind people when collaborating through virtual environments. We hope to achieve this by developing more collaborative applications and testing the awareness mechanisms we use. Especially, we have to investigate about the feeling users have about what is going on in the collaborative game and how do they use the information provided.

Acknowledgment

This report was funded by the Chilean National Fund of Science and Technology, Fondecyt, Project 1030158.

References

1. Borges, M., Pino, J. (1999).: Awareness mechanisms for coordination in asynchronous CSCW. Proc. Of the WITS'99 Conf. Charlotte, North Carolina, Dec 1999, pp. 69-74
2. Cohen, J. (1994).: Monitoring background activities. In G. Kramer (Ed.), Auditory display: Sonification, audification, and auditory interfaces. Reading, MA: Addison-Wesley, pp. 499-522
3. Collazos, C. Guerrero, L. Pino, J. Ochoa, S. (2002): Introducing shared-knowledge Awareness. Proceedings of the IASTED Conference Information and knowledge sharing. Nov. 2002, St Thomas, Virgin Islands, pp. 13-18
4. Dourish, P. Belloti V. (1992): Awareness and coordination in shared workspaces. Proceedings of CSCW'92, pp. 107-114
5. Gross, T., Prinz, W. (2003): Awareness in Context: A Light-Weight Approach, Proceedings of the European Conference on Computer-Supported Cooperative Work (E-CSCW), Helsinki, Finland, Sept. 2003, pp. 295-314
6. Gutwin, C. and S. Greenberg (1996): Workspace Awareness for Groupware. Proceedings of the Conference on Human Factors in Computing Systems. Vancouver, pp. 208–209
7. Isaacs, E., Walendowski, A. and Ranganathan, D.: (2002) Hubbub: A sound-enhanced mobile instant messenger that supports awareness and opportunistic interactions. Proceedings of the Conference Computer-Human Interaction (CHI), Minneapolis, MN, April, 2002, pp. 179-186
8. Lumbreras, M., & Sánchez, J. (1999): Interactive 3D Sound Hyperstories for Blind Children. Proceedings of the ACM CHI '99, Pittsburg PA, USA (1999), pp. 318-325
9. Moran, T., Anderson, J. (1990): The workaday World as a paradigm for CSCW Design. Proceedings of CSCW '90. October 1990, pp. 381-393
10. Morley, S., Petrie, H., O'Neill A.-M. & McNally, P. (1998): The use of non-speech sounds in a hypermedia interface for blind users, in Edwards, A.D.N., Arato, A., and Zagler, W.L.(Eds.)'Computers and Assistive Technology'. Proceedings of ICCHP'98. XV, IFIP World Computer Congress, pp. 205-214
11. Ressler, S. & Antonishek, B. (2001): Integrating Active Tangible Devices with a Synthetic Environment for Collaborative Engineering. Proceedings of the 2001 Web3D Symposium. Paderborn, Germany, February 19-22, pp. 93-100
12. Roth, P., Petrucci, L., Assimacopoulos, A. & Pun, Th. (2000): Concentration Game, an Audio Adaptation for the blind. CSUN 2000 Conference Proceedings, Los Angeles, USA, pp. 1-4
13. Sánchez, J. (2002). VirtualAurea: Perception through spatialized sound. Proceedings of the 4th International Conference on Disability, Virtual Reality and Associated Technologies, September 18-20, 2002,Veszprém, Hungary, pp. 67-73
14. Sánchez, J., Baloian, N. & Hassler, T. (2004). Blind to Sighted Children Interaction through Collaborative Environments. Proceedings of the CRIWG ´04. Lecture Notes in Computer Science 3198, Springer-Verlag, pp. 192-205
15. Sohlenkamp, M. Supporting group awareness in Multi-user environments through perceptualization. GMD research series No. 6 (1999), 172 pp

Modeling 3D Interactive Environments for Learners with Visual Disabilities

J. Sánchez and N. Baloian

Department of Computer Science, University of Chile
Blanco Encalada 2120, Santiago, Chile
{jsanchez, nbaloian}@dcc.uchile.cl

Abstract. Educational software has been criticized for not using explicit models to generalize and replicate good practices. Actually almost every educational program has a model, but most of them remain implicit. In this paper we propose a methodology for developing educational software for children with visual disabilities. Multimedia software for these children has some particularities reflected on our model with emphasis on process modelling including learner evaluation and feedback. The model emerges from research on developing educational software for children with visual disabilities and studies concerning the design of educational software for sighted learners. The model was validated by special education teachers and software designers trying the model with five software products based on model heuristics.

1 Introduction

The design of educational software should make extensive use of all multimedia capacities modern computers can offer. This implies that users with visual disabilities are unable to access to this software because of the use of graphical user interfaces.

Although not as much as for sighted people, there have been some developments of educational software for people with visual disabilities. However, they usually lack of critical interface elements commonly present in software for sighted children. Most software does not include explicit model knowledge and skills learners should construct when using the software, an explicit learner model, and the implementation of appropriate feedback to improve the learners' performance. To many authors, designers of educational software for children with disabilities conceive the software with interaction restrictions in their minds, fixing the interaction modes from the very beginning. Thus software is from the beginning conceived with limitations. We propose that educational software for learners with visual disabilities should be designed without taking into account from the beginning the users' disabilities. They should start by considering relevant modeling aspects. Only when it comes to the point of mapping the inputs and outputs of models into an interface, the learner capabilities and disabilities should be taken into consideration to map these variables on proper devices.

Since educational software development process depends on people, tools, and methodologies involved, and considering that we have not a clear methodology to carry out this process for children with visual disabilities, the results mainly depends on the skills

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 1326 – 1333, 2006.

of the involved people. This can cause many drawbacks typical for a hand-crafted process. Software engineering uses methodologies to help to reduce the craftsmanship level of software development by using the best methodological practices.

There have been some proposals for methodologies to develop educational software (Alessi & Trollip, 2001; Soares, 2001; Dillenbourg & Self, 1992) and courseware (Baloian et al., 2001). An interesting question is whether we need another new methodology for this kind of software or if the existing one can work appropriately. We think the process should start from existing, well accepted methodologies extending and adapting them to this particular case in order to guide the implementation team to think about the possibility of including elements of intelligent tutoring systems in software aimed to blind users. The goal of this methodology is to assist developers in considering critical components for educational software design. The methodology proposes also a characteristically system architecture.

Software based on virtual environments for users with visual disabilities are based on the presentation of graphic information by text-to-speech translation that reads Web pages displayed through browser and three-dimensional spaces of navigation environments with sounds that can get close, far or move to mentally represent the space (Mereu & Kazman, 1996) and to develop cognitive skills (Savidis et al., 1996). This can be seen in Morley et al. (1998), where blind people develop a special way of navigating through a known environment and represent spatial structure with cognitive difficultness. The system uses different output devices such as concept keyboard, tablets, switches, tactile interfaces (Lange, 1999), and forcefeedback (Ressler & Antonishek, 2001).

The HOMER UIMS was produced by Savidis and Stephanidis (1995), Savidis et al., (1996) developing dual interfaces to integrate blind and sighted learners. HOMER integrates visual and non visual interaction with objects and their relationships. The browser BrookesTalk reproduces a Web page by using synthesized voice with words, sentences, paragraphs, and offering different points of view of the page to simulate scanning (Zajicek et al., 1998).

A game for audio concentration is presented by Roth et al., (2000) consisting of pairing different levels of geometric figures, basic, and derives. To represent geometric figures graphically they constructed a bi-dimensional sound space. This concept allows graphic representation such as icons to be represented by the perception of moving sounds in the spatial plane. Blattner and Brewster introduced "earcons" as non verbal audio messages to provide information to users about computer objects, operation, and interactions (Blattner et al., 1998; Brewster, 1998). Each dimension corresponds to a musical instrument and the points of the plot correspond to pairs of frequency in a scale. The horizontal movements from left to right are equivalent to a frequency variation of the first instrument and the vertical movement to frequency variations of the second one.

AudioDoom (Lumbreras & Sánchez, 1999) allows blind children to explore and interact with virtual worlds by using spatial sound. The game was based on the traditional Doom game where the player moves through corridors discovering the environment and solving problems simulated with objects and entities that inhabit a virtual world. VirtualAurea (Sánchez, 2001, 2002) was developed after it was proved that sound-based virtual environments can help to develop tempo-spatial cognitive structures of blind children. It is a spatial sound tool editor that can be used by parents and teachers to design a wide variety of spatial maps. Users can integrate different sounds by associating them to objects and entities in a story. AudioMemory (Sánchez & Flores, 2004) a virtual environment based on audio to develop and use short-term memory. It was also modeled with mathematics contents, AudioMath, to assist the learning of basic mathematics operations such as multiplication and division. Results evidenced that both software helped to develop and enhance memory and mathematics learning in blind children. AudioChile (Sánchez & Sáenz, 2005) is a 3D interactive environment for children with visual disabilities to help them to solve problems related with the Chilean geography and culture. AudioChile can be navigated through 3D sound to enhance spatiality and immersion throughout the environment. 3D sound is used to orientate, avoid obstacles, and identify the position of diverse personages and objects within the environment.

2 Process Modeling

The methodology proposed is based on the following hypotheses: 1. The knowledge and skills a learner has to develop with the aid of the software are measurable and can be represented. This implies that a learner's model can be constructed, and 2. The software represents an interactive environment (real or imagined) for learning. This means that the software allows the construction of knowledge by the learner

The "workflow" for developing educational software for learners with visual disabilities proposed by this methodology is depicted in the Figure 1. To explain this methodology we use AudioDoom (already introduced in chapter 1). Normally, a software development process starts with the definition of software requirements. In this case, the requirements are represented by the learning goal that in the case of AudioDoom is the ability to create a mental model of the surrounding environment. According to the learning goal, an appropriate scenario should be conceived to allow learners to develop a skill or knowledge. In AudioDoom the idea was having the learner to discover a labyrinth full of sound emitting objects and entities. The next step is modeling the environment. At the same time, and based on artificial intelligence strategies, the learner's knowledge should be modeled (Baloian et al., 2002). Developing and describing a model has its own process (Zeigler, 1976). The result of this step is a formal model description for both, the learning environment and the learner's knowledge in paper. A computer program has to implement this. At this point, it is important to consider the development of an (or use an already existing) editor for generating different environments, such as editors for constructing different labyrinths, instead of having a single environment "hardwired" represented by the program. After this process, model input and output variables are clearly identified. Then we need to map or project them on input/output devices suitable for children with visual disabilities.

As we see in Figure 1, cognitive goals will not only influence the definition of learning environments but also the generation of metrics for evaluating knowledge construction by the learner. This will be discussed in more detail in the following chapter. Then learners should explore the environment as a way of testing. Test results should be evaluated through usability methods and determine the effectiveness and impact to help learners to achieve the cognitive goals. The evaluation may cause a revision of the real world representation, the model, and the interface. Revisions of the real world and its modeling can be mostly caused by the failure of the software's effectiveness in supporting learners to achieve the cognitive goals: the environment does not provide the adequate learning activities and the model does not implement them properly. Revisions of the interface (projection of the input/output values on adequate devices) can be caused by usability drawbacks in the software.

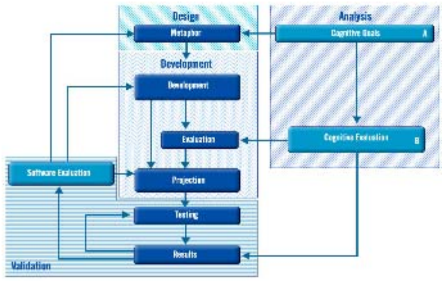


Fig. 1. The workflow for developing educational software for learners with disabilities

Analysis: Consists of two sub-stages, A and B. The first stage is to define cognitive goals to be achieved by the learner. This corresponds to the definition of software requirements. The second stage is to define procedures and functions to evaluate the achievement of cognitive goals. *Design***:** In this stage a metaphor will be defined for a "world" or scenario where the learner constructs knowledge through the interaction with this world. Normally, this is game type of software and "playing rules" are defined. This leads to define the model of the world and the knowledge to be constructed. *Development:* Consists of three sub-processes. The first process is the computational implementation of models of the world and learner. We recommend exploring the possibilities of developing editors for implementing different scenarios of the same "world". The second stage is the implementation of the evaluation process and the feedback to the student. The third stage is the projection of the models. We identify input and output variables of models as well as parameters and results (including the feedback to the learner) of the evaluation function. These values have to be "projected" properly over the haptic, audio, and visual (for people with residual vision) input/output devices available. We verify a wide variety of input/output devices to avoid limiting to traditional devices such as joysticks and keywords for input and sound for output. Haptic devices such as tablets, electronic boards, and Phantom can give blind users sensations of being "touching" virtual objects. It is important to make these actions after setting the models in order to ovoid restricting from the beginning of software design. Some guidelines to implement the projection are given below based on the literature and our own experience in developing software for blind children. *Validation:* Consists of two sub-processes. First, we develop usability tests to get data about how well the system fit our objectives in order to attain the cognitive goals set at the beginning. We emphasize the analysis of some elements of human-computer interaction. Second, we analyze these results and study how the metaphor, models, and the projection of input/output variables can be improved. Normally an error in the integrity of the system for learning can imply to review the metaphor and models used. Usability issues can lead to review the projection.

3 Modeling the Resulting Architecture

Fig.2. represents the architecture of the resulting software containing the following components:

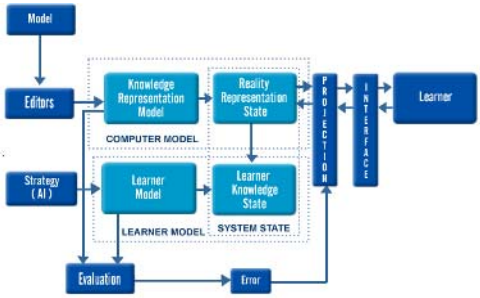


Fig. 2. Architecture of the resulting software

Metaphor of the real world (Model): According to the cognitive skills to be developed real world metaphors are designed as well as the activities learners have to do to attain the cognitive goals by considering their interest and motivation. *Editor:* Tools to construct an internal model based on 2D/3D graphic representations or auditory representations. *Computer representation of the real system:* Corresponds to the computational representation of the problem, the real world metaphor, and models the knowledge. Contains functions, parameters, and variables of the state of the system describing the situation of the represented world and how the transition from one state to other will be made by considering the interaction of the learner with the software and reflected on the entry variables. *Strategies*: This component gives the strategies to be used to model the state of the learner's knowledge. They are taken from the field of artificial intelligence applied to intelligent tutoring. Some of them are the overlay model (Kass, 1989) that treats the learner's knowledge as a subset of an expert knowledge. The Differential model (Clancey, 1987) extends the previous model by dividing the learner's knowledge in two categories: knowledge that the learner should know and knowledge that is not expected to be known by learners. The perturbation model (Kass, 1989) supposes learner should posses a potentially different knowledge in quantity with respect to an expert. The model can represent the knowledge and beliefs of the learner beyond the ranks of the expert's model. *Learner model*: This component represents what the system thinks about the state of the students learning in a certain point. It contains knowledge and skill representations the learner should construct, the variables of the state of learner representing the level of learning in a certain moment, and the rules about how to upgrade this information given the interaction with the system and reflected in the change from one real world model state to another. Thus the learner model is given by making inference of the individual knowledge by analyzing the performance (Dillenbourg & Self, 1992). *Evaluation:* This component defines the difference between the knowledge model represented in the software and the knowledge model of the learner generated by the strategies. Thus an error measure is produced and projected to the interface as student's feedback. *System* *projection*: This is the main component to certify that the software can be fully assimilated by children with visual disabilities. It is in charge of projecting most interactions, state variables, and feedbacks from and to the software.

4 Model Evaluation

We tested our model with three special education teachers and two software developers. They evaluated the model by analyzing five products to check how well they meet the methodology proposed above. To do this we designed a Likert type scale based on model heuristics. From the model proposed above we defined four major heuristics for evaluation purposes: metaphor, learning, interaction, and interface. Metaphor included adequacy to the mode of learning, how well it represents the model, and if it defines different interaction environments (editors). Learning includes if the software represents what learners have to learn, evaluates learning adequately, and provides feedback to the learner. Interaction includes if the input/output devices are adequate, and how well users can orient and know what to do and where to go by themselves. Interface included font and typography, colors, buttons, icons, audio cues, and feedback used.

The results of the model evaluation are presented in Table 1. Possible answers spanned from "do not meet the heuristic" (1) to "highly meet the heuristic" (5). Average resulting scores were from 3.4 (VirtualAurea) to 4.6 (AudioMemorice), evidencing that most software analyzed meet the minimum standards posed by the model.

Table 1. Model evaluation results. 1. AudioBattleShip, 2. AudioMemorice, 3. VirtualAurea, 4. Theo&Set, 5. CantaLetras, 6. AudioVida.

Indicators / Software	1	2	3	4	5	6	Heuristics
The metaphor is adequate for the learning method (construction)	4,8	4,8	4,0	4,5	4,7	3,9	METAPHOR
The model represents well the metaphor	4,5	4,8	3,5	4,8	4,7	4,2	
It is possible to define different interaction environments (editors)	5,0	4,8	5,0	4,3	3,7	3,0	
The software represents somehow what learners have to learn	3,8	4,5	4,3	4,8	4,8	4,0	LEARNING
The software evaluates coherently what the learner have to learn	3,3	4,0	3,5	4,3	5,0	3,7	
The software provides adequate feedback to the learner	4,5	4,8	3,5	4,3	4,5	3,0	
There are input/output devices for interaction purposes	4,3	5,0	4,3	4,5	3,8	3,7	INTERACTION
Users can know where they are	4,3	4,0	3,8	3,8	2,5	2,8	
Users can know what to do in any moment	3,5	3,8	3,3	3,5	2,3	2,9	
The font typography used is adequate	1,3	4,7	1,0	4,3	2,5	4,1	INTERFACE
The size of the font is adequate	1,3	4,7	1,0	3,8	2,0	4,0	
Colors and contrasts are used adequately	4,6	4,8	2,3	4,5	3,5	3,4	
The design of buttons and icons are adequate	2,5	4,8	1,5	3,8	2,5	3,6	
The interface generates adequate audio feedback for learners with visual disabilities	4,3	4,5	3,5	3,5	4,0	3,4	

From the results displayed we can state four initial conclusions. First, we validated the model by evidencing that using heuristics is a clear methodology for model analysis in educational software. Second, all products considered the heuristics in different degrees. Third, metaphor and learning are the heuristics that best meet the standards of our model. Fourth, interaction and interface were the least attained heuristics. Then our model was initially validated with existing educational software through walk-through techniques used by teachers and software developers.

5 Discussion and Further Work

We present a methodology for developing educational software for children with visual disabilities. The model is the result of a growing need for models to develop, replicate, evaluate, and improve educational software for this population. This is a process model with the resulting architecture including ways of evaluating and giving feedback to learners, as well as to set qualitative differences for children with and without visual disabilities. We describe formally and operationally the model and propose some guidelines to design educational software for children with visual disabilities by discussing main generic attributes to include in this software.

The model was tested for viability in educational software design. Interesting results came out when teachers and software developers went through existing educational software for blind children by using some heuristics drawn from the model. Most software did meet the heuristics in different degree. Interactivity and interfaces were the least ranked, meaning that these heuristics need to be carefully considered when design software for children with visual disabilities. Now we need first to improve our heuristics and evaluation instruments, and then apply them to different learning contexts for children with visual disabilities. The next step will be to design and develop software for children with visual disabilities by following step by step the methodology proposed here. Finally, we expect to contribute to the field with an explicit and functional model that can be generalized and replicated to help to improve the learning of children with visual disabilities.

Acknowledgements

This report was funded by the Chilean National Fund of Science and Technology, Fondecyt, Project 1030158.

References

1. Alessi, S. & Trollip, S. (2001). Multimedia for Learning, Methods and Development. Boston: Allyn and Bacon
2. Baloian, N., Ochoa, S. & Fuller, D. (2001). A model for component-based courseware development. Proceedings of the KSI Conference SEKE'01. Buenos Aires, Argentina, Jun. 13-15. KSI Press. USA pp. 363-370
3. Baloian, N., Luther, W., Sánchez, J. (2002). Modeling Educational Software for People with Disabilities: Theory and Practice. Proceedings of the ACM ASSETS 2002 Conference, pp. 111-117
4. Blattner, M., Sumikawa, D. & Greenberg, R. (1998). Earcons and icons: Their structure and common design principles. Human-Computer Interaction, 4, (1) (1998), pp. 11-44
5. Brewster, S.A (1998). Using earcons to improve the usability of a graphics package. Proceedings of HCI'98, Sept. 1-4(1998) Sheffield Hallam University, Sheffield, UK
6. Clancey, W.J.(1987). Methodology for Building an Intelligent Tutoring System. En Kearsley, G., Artificial Intelligence and Instruction - Applications and Methods, pp.193-227

7. Dillenbourg, P. & Self, J. (1992). A Framework for Learner Modeling. Interactive Learning Environments, 2(2), pp. 111-137, 1992
8. Kass, R.(1989). Student Modeling in Intelligent Tutoring Systems –Implications for User Modeling. In A. Kobsa & W. Wahlster, Eds. User Models in Dialog Systems, pp. 386-410.
9. Lange, Max O. (1999). Tactile Graphics – as easy as that. CSUN's 1999 Conference Proceedings. March 15-20, Los Angeles, USA
10. Lumbreras, M., & Sánchez, J. (1999). Interactive 3D Sound Hyperstories for Blind Children. Proceedings of the ACM CHI '99, Pittsburg PA, USA (1999), pp. 318-325
11. Mereu, S.& Kazman, R. (1996). Audio enhanced interfaces for visually impaired users. Proceedings of the ACM CHI 96, pp. 72-78
12. Morley, S., Petrie, H., O'Neill A.-M. & McNally, P. (1998). The use of non-speech sounds in a hypermedia interface for blind users, in Edwards, A.D.N., Arato, A., and Zagler, W.L.(Eds.)'Computers and Assistive Technology'. Proceedings of ICCHP'98. XV, IFIP World Computer Congress, pp. 205-214
13. Ressler, S. & Antonishek, B. (2001). Integrating Active Tangible Devices with a Synthetic Environment for Collaborative Engineering. Proceedings of the 2001 Web3D Symposium. Paderborn, Germany, February 19-22, pp. 93-100
14. Roth, P., Petrucci, L., Assimacopoulos, A. & Pun, Th. (2000). Concentration Game, an Audio Adaptation for the blind. CSUN 2000 Conference Proceedings, Los Angeles, USA.
15. Sánchez, J. (2001). Interactive virtual acoustic environments for blind children. Proceedings of ACM CHI ´2001, Seattle, Washington, April 2-5, 2001, pp. 23-25
16. Sánchez, J. (2002). VirtualAurea: Perception through spatialized sound. Proceedings of ICDVRAT, pp. 67-73. September 18-20, 2002, Veszprém, Hungary
17. Sánchez, J., Baloian, N. & Hassler, T. (2004). Blind to Sighted Children Interaction through Collaborative Environments. Lecture Notes in Computer Science 3198, Springer-Verlag, pp. 192-205
18. Sánchez, J.,Flores, H. (2004). Memory enhancement through audio. ACM SIGACCESS Accessibility and Computing Archive Issue 77-78, pp. 24-31, Sept. 2003-Jan. 2004
19. Sánchez, J., Sáenz, M. (2005). 3D Sound Interactive Environments for Problem Solving. Proceedings of The Seventh International ACM SIGACCESS Conference on Computers and Accessibility, Assets 2005, Baltimore, Maryland, USA, October 9-12, pp. 173-178
20. Sánchez, J., Baloian, N. (2005). Modeling Audio-Based Virtual Environments for Children with Visual Disabilities. Proceedings of the World Conference on Educational Multimedia, Hypermedia & Telecommunications ED-MEDIA 2005. Montreal, Canadá. June 27 July 2, 2005, pp.1652-1659
21. Savidis, A. & Sthphanidis, C. (1995). Developing dual user interfaces for integrating blind and sighted users: HOMER UIMS, ACM CHI'95 Conference Proceedings, Denver, CO (1995), pp. 106-113
22. Savidis, A., Sthphanidis, C., Korte, A., Crispie, K. & Fellbaum, K.(1996). A generic direct-manipulation 3D-auditory environment for hierarchical navigation in non-visual interaction. Proceedings of the ACM ASSETS 96, pp. 117-123
23. Soares, M.P.S. (2001). Modeling educational software for teaching quality control with emphasis on quality evaluation software, MAQS. Proceedings of the International Symposium on Computing in Education, Aguas Calientes, México
24. Zajicek M., Powell C. & Reeves C. (1998). A web navigation tool for the blind. Proceedings of the ACM ASSETS'98, 3rd ACM/SIFCAPH Conference. on Assistive Technologies, Los Angeles, USA (1998), pp. 204-206
25. Zeigler, B. P. (1976). Theory of Modeling and Simulation, Wiley, New York, USA, 1976

APL: Audio Programming Language for Blind Learners

Jaime Sánchez and Fernando Aguayo

Department of Computer Science, University of Chile
Blanco Encalada 2120, Santiago, Chile
{jsanchez, faguayo}@dcc.uchile.cl

Abstract. Programming skills are strongly emphasized in computer science. Programming languages are constructed based on sighted people as end-users. We have designed Audio Programming Language for blind learners based on audio interfaces to support novice blind learners to develop and exercise problem solving skills. APL was designed with blind learners from the beginning to construct programs and solve problems with increasingly complexity. Audio Programming Language was usability tested during and after implementation. Blind learners used, wrote programs, and helped to make improvements to this programming language. Testing results evidence that APL mapped the mental models of blind learners and helped to motivate them to write programs and thus entering to the programming field.

1 Introduction

Diverse attempts have been made to make programming closer to end-users: Basic, Logo, Smalltalk, Pascal, Boxer, Playground, KidSim, AgenSheets, LiveWord, Shoptalk. All of these programming languages have contributed to expand the number of people who can program. Many of them have applied user interfaces principles to programming. This has ended up with better skills for novice learners to program. Thus the literature describes studies concerning programming by demonstration, programming by example, visual programming, graphical programming, and physical programming [2,3,5,7]. Most of these programming attempts (if not all) have been focused on visual programmers.

To write a program using these languages learners with visual disabilities has to use text-to-speech systems that "read" programming commands and variables assuming that they follow easily the same logic of programming used by sighted learners.

Recent studies have shown that by using audio-based applications blind children can develop and rehearse cognition [1,4,6,8,9,10]. Most of these studies focus on the development of 3D audio interfaces to map the entire surrounding space and thus helping blind children to construct cognition through audio-based interfaces such as tempo-spatial relationships, short-term memory, abstract memory, spatial abstraction, haptic perception, and mathematic reasoning.

This research study introduces APL, Audio Programming Language. APL is based on audio to enhance problem solving and thinking skills in novice blind learners. APL is not thought as an alternative to conventional programming languages, rather it is a

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 1334 – 1341, 2006.

tool to help to better integrate novice programmers to the field of conventional languages. In doing so APL can generate Java code and thus make that novice programs can be used by other people without needing APL. They will just need Java and a TTS engine. APL is a tool to motivate blind learners to enter to the programming world.

2 Programming by Blind Learners

What are the specific needs of blind learners to program? Why is it difficult for novice blind learners to map and follow current programming languages? Can we develop a language to fit the needs of these learners, especially those that will not follow a computer science career? These were some of the underlying questions in our study.

A programming language communicates the programmer with the computer. To do this the structure and logic are designed in such a way to be interpreted by the machine. Actual languages are based on the idea that a programmer writes command lines interpreted by the computer. These commands must be correctly written and well defined otherwise the machine cannot understand instructions and specific tasks are unsolved. This implies to memorize a huge amount of command lines and to write them correctly in order to avoid error parsing. Most of these programming languages are heavily based on visual interfaces.

Two major difficulties can be found when blind learners use these languages. First, if they use a pure language they face the issue of verifying the program consistency and the correct reading of command lines. Second, if we provide them current tools to support program construction, they deal with graphical user interfaces. APL intends to close the gap between programming by sighted learners and programming by blind learners.

Fig. 1. A blind learner programming with APL

APL is the first intent to develop a programming language oriented to blind learners. They can interact and communicate with the computer as sighted programmers do by writing their own programs and exchanging them with other users, sighted or blind. The only difference is the way they do it and the sensory channel used to interact with the machine.

To attain our goals APL was designed to make easier the correct writing of command lines and to avoid parsing problems. We facilitate machine-programmer interaction through additional tools not commonly found in current programming languages.

Common programming languages use minimum storage units or variables that can be appropriate to the user's needs. It is hard to imagine a language without string or integer variables. In contrast, APL implements a new and unconventional type of variable to store sounds to be later manipulated by blind users. Sound is a fundamental unit of APL.

3 Audio Programming Language

We aimed at developing a tool to facilitate the manipulation of a programming language by blind learners. APL is oriented to blind novice programmers by diminishing syntax complexity of a current language and adding functions to analyze the final program. Learners can get programs in Java code with the help of an APL code generator that makes the necessary conversions to add the required functionalities.

3.1 Model

We developed an implementation model for APL. Two basic conditions were set: The software should be easy to implement and flexible by allowing small changes without having to modify the whole program. APL has three main layers: Audio interface, programming interface and programming logic (see Figure 2).

Audio Interface includes two states: TTS and Recorded Text. TTS is used to reproduce words and sentences. Recorded Text is used to give command feedback when running the round list and to pronounce letters and numbers.

Programming interface contains three states: Round list, keyboard, and Menu. Round list includes all commands used in the program since the list consistency is operated dynamically by the command integrity. Keyboard, which can be activated any time the novice programmer enters information to the program such selecting a command from the round list, variable names, and text for output. Menu is active during the whole use of APL and allows executing the program or exporting this to Java code

Programming Logic includes four states: Command integrity, free input, run program, and export to Java. Command integrity is in charge of maintaining the integrity of the program by including or eliminating commands in the round list to help programmers to use adequate commands and communicating with the TTS or recorded text. Free input allows users to enter text freely providing permanent feedback through the integration of the TTS (words and sentence) and the recorder text (letters and numbers). Run program executes the program if the command state is correct. Export to Java exports the program through a code generator.

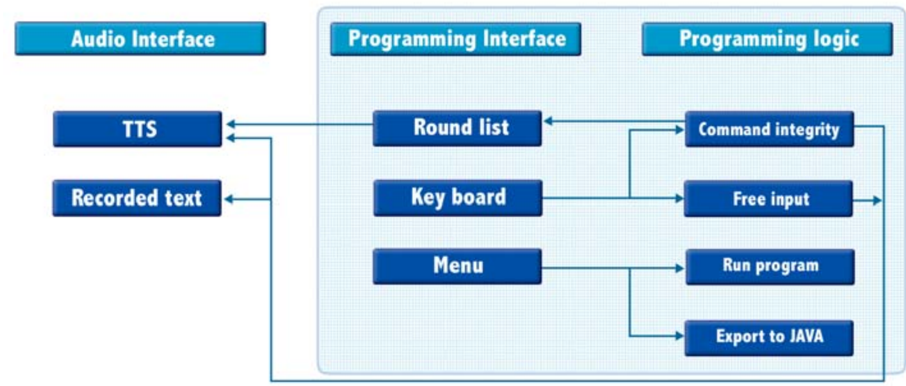


Fig. 2. Layers of APL

3.2 Implementation

APL was developed by using Java 1.4.1 and FreeTTS Java speech synthesizer. These are fine tools for media management and close to the machine logic. APL has the following modules: DataBase, Integrity, Kernel, CHI and generador de JAVA.

DataBase stores the commands and its integrity. Integrity reads the DataBase and loads a certain structure and commands. Kernel manages the interaction between the user and programming rules. CHI controls input/output audio and translates input/output text to audio. The interface is entirely based on audio in such a way that blind programmers have not direct interaction with the screen. APL has two modes: programmer mode and running mode.

There is no programming language skills prerequisite to use APL. Learners don't need to memorize any commands. APL has a circular command list introduced to the user to select the desired command. The list has options related to the actual state of the system. In doing so two current issues in conventional programming can be solved: command navigation and command semantic.

To control command navigation APL presents a reduced command list to improve the navigation through the list and to optimize the programming time. Command semantic is prevented by assuring that the selected command is correctly written in such a way that APL can interpret it in the runtime.

The circular list of APL is dynamic. The main list consists of input, output, cycle, condition, and variable (see Figure 3). Each of these commands possesses an integrity table placed on the programming logic layer (see Figure 2) to upgrade the circular list with new commands. APL can request answers to a set of questions to complete the command line at the audio interface layer. This process can be repeated as many times as necessary to write a program. The command definition is:

Variable: It consists of variable definitions. They can be text, numeric or sound. Text is read through text-to-speech.

Input: The blind programmer uses this command to create an input requirement when executing the program by selecting the desired variable, keyboard or sound. This input should be saved as a variable defined by the programmer.

Output: It is the medium available to the programmer to create an output of the product. Here APL uses only sound interfaces. Sound has two modes: 1. The programmer defines a variable and saves it with a corresponding sound, and 2. The programmer defines a variable and saves it with a corresponding word or sentence entered through the keyboard and read as text-to-speech when executing the program.

Condition: It is a command that compares variables defined by the programmer. If variables are true the commands between tags are executed and a new option is created in the command list, end of condition.

Cycle: It is a command that allows executing the program many times in a defined sector. It comes out of the sector when the defined cycle condition is attained. The cycle is completed by using the end cycle option.

For feedback optimization when the programmer writes the program we used a mix solution of TTS and recorded text to provide an immediate response when writing since we used recorded letters and numbers to avoid TTS processing.

4 Usability Testing

APL was built by and for blind learners. We implemented a qualitative case study methodology by observing and recording the interaction of end-users and experts with the program. APL was usability tested during and after development.

Expert users: Immediately after APL was implemented we tested the software with three expert users that have experience and knowledge in programming to validate the functionality and to identify possible problems and errors in order to get feedback, suggestions, and make some improvements. They were informed about the correspondence of APL commands with generic commands and wrote programs using APL. They also made comments and answered pre-designed questions detecting functioning problems such as: a. APL did not support multiple cycles, b. APL did not allow to make waterfalls conditions, c. The conditions of APL did not allow to exit from cycles, d. Some issues occurred when manipulating "keyboard" variables (variable types such int, var, and string cannot be defined), e. Sound and voices can be confusing in some passages, f. The executor went down when the cycle was not appropriately defined. Then we analyzed in depth the content and comments given by experts and redesigned the prototype before the blind users tried it.

End-users: APL was evaluated through a qualitative case study methodology by observing and recording the user's interaction with APL. The study included a two stages usability pilot testing with three blind novice programming learners, ages 17 to 20. During the first stage learners followed four interactive sessions of 2.5 hours each. First, basic programming concepts were introduced to them and after that they solved related mental tasks. Second, they interacted with APL by using basic commands to create simple programs such as Hello World and guessing games. Third, they wrote their own programs such as "think an idea and make a program". Finally, they wrote a small program for blind and sighted people to learn Braille (see Figure 1).

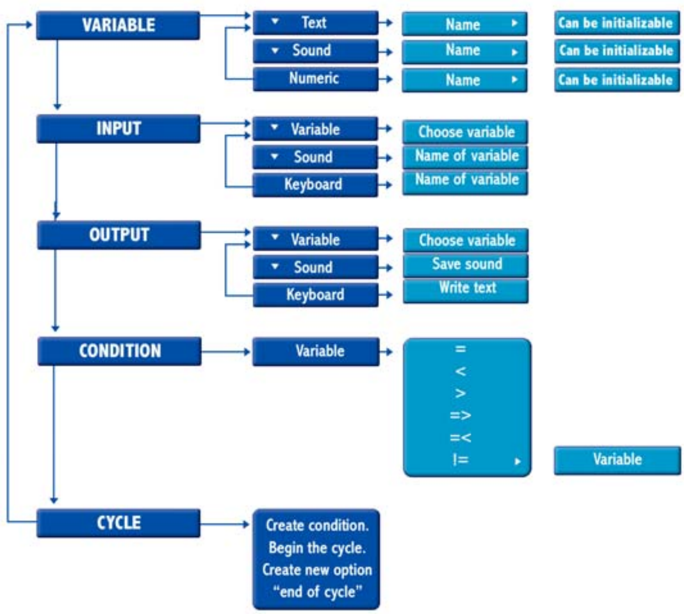


Fig. 3. Command list tree

The main goal was to familiarize users with concepts and elements that can be used to solve problems with APL. The idea was having users to interact with APL, to identify classic algorithms and data structure in computer science, and to develop algorithmic thinking to solve a problem. Users understood the functioning and the main goal of APL. They also learned about the basics of programming.

The second stage consisted of answering a usability evaluation test and having an interview. The test consisted of eighteen user satisfaction questions in a Likert type scale. The interview pursued to get in depth answers about APL and to record comments about their experience. Blind users' behaviors were observed by a team of three researchers.

In the beginning learners understood some programming concepts but it was complex to them to understand the concept and functions of different commands such as cycle, condition, and variable. This understanding took more time as expected. During the third session they understood these concepts and made comments concerning APL based on their interaction experience and expectations about the functionality. We observed that methodologies such as theoretical explanations and step by step exercises did not help them to understand fully the basic concepts of programming. This can be explained due to the fact that blind learners tend to rely heavily on concrete experience before building abstract thinking. They did not construct a mental model of APL and programming until they designed their own idea and understood the meaning fully. Then through passing from the known to the unknown they increasingly developed understanding of programming.

Blind learners were able to interact with APL by using commands such as input, output, variables, cycles, and condition. They could write some programs following step by step directions as well as writing programs emerging from their own ideas. During the last session they evidenced to understand programming concepts and apply them to their programs. In doing so blind users constructed logic thinking skills, verbalized them with their partners and observers, and wrote programs to prove their feasibility. They were very motivated and satisfied as a result of interacting with APL. We assume these are preliminary results but very promising.

As users understood concepts, used commands, and applied them to their programming they were able to design logic instructions, understand, reproduce, and verbalize them. They also constructed algorithms that could be reproduced mentally. They showed high interest and enthusiasm when programming.

Blind learners also made diverse comments and suggestions to improve APL such as to include a feedback by using the space bar to remind the last written word, to start up APL by their own, and to make it faster when using arrows.

Our experience with blind users programming APL shows that current programming visual tools are not feasible for these learners. It appears that the issue for novice learners is not just to adapt visual programming to blind learners [11]. The theoretical background of these tools is complex to be easily mapped by novice blind learners. One interesting result of our usability test evidenced that interactivity may have rather different meaning and implications to blind users.

TTS and recorded text were tested separately from the use of APL to evaluate TTS and recorded text feedback for letters and number pronunciation. Results were entirely satisfactory demonstrating the user's preference for the integrated system (TTS and recorded text) instead of the previous system with a TTS processing all letter conventions.

5 Conclusion

We have introduced APL, Audio Programming Language for blind learners. APL is designed to help novice blind learners to enter to the programming world and to solve problems and develop thinking skills by targeting their needs and mental models. Learners were able to interact and program APL by using audio-based commands. They understood programming concepts and apply them by programming their own ideas.

Adequate uses of audio as a sensory interactive medium can help learners to learn how to program. We initially observed that blind learners can develop algorithmic and logic thinking skills with APL. They verbalized these skills, and wrote programs to prove their feasibility. The experience of interacting with APL was motivating and unique to them by expressing satisfaction with the programming experience. They have continued to program with APL after participating in the study.

APL is not an alternative proposal to visual programming for blind users. We believe that for higher level tasks there should be interfaces to allow an adequate access to visual programming such as in [12]. Our proposal goes in a different direction. We aimed at motivating novice blind learners to learn programming basic ideas and solve problems with APL. Our idea is that their programs can be used by other people. To do this we have embedded in APL a Java code generator.

We are learning the way blind users map the programming process by using APL. It may be somehow different from sighted users. Interacting through programming may have a different meaning to them. This should be studied more fully in future studies. Finally, APL is a first step to provide a robust Audio Programming Language to blind learners and thus helping them to enjoy programming and developing their cognition.

Acknowledgments

This report was funded by the Chilean National Fund of Science and Technology, Fondecyt, Project 1030158.

References

1. Baldis, J. Effects of spatial audio on memory, comprehension, and preference during desktop conferences. Proceeding of the ACM CHI ´01, Vol 3, 1, (2001), pp. 166-173
2. Cypher, A., Halbert, D. C., Kurlander, D.: Lieberman, H., Maulsby, D., Mayers, B.A., and Turransky, A. (eds.). Watch What I Do: Programming by Demonstration. Cambridge, MA: The MIT Press, 1993.
3. Lieberman, H. Your Wish Is My Command: Programming by Example. Morgan Kaufmann, San Francisco, 2001.
4. McCrindle, R. & Symons, D. Audio space invaders. Proceedings of the Third International Conference on Disability, Virtual Reality and Associated Technologies, (2000), pp. 59-65.
5. McDaniel, R. and Myers, B. Getting more out of Programming-By-Demonstration. In Proceedings of CHI ´99. ACM, Pittsburgh PA, 1999, pp. 442-449
6. Mereu, S. & R. Kazman. Audio enhanced 3D interfaces for visually impaired users. Proceedings of CHI '96, pp. 72-78 ACM Press. (1996)
7. Montemayor, J. Physical programming: software you can touch. Proceedings of ACM SIGCHI, ACM Press, pp. 81-82, March 2001
8. Sánchez, J. Interactive 3D sound hyperstories for blind children. Proceedings of ACM CHI ´99, (1999). Pittsburg PA, pp. 318-325
9. Sánchez, J. AudioBattleShip: Blind learners' collaboration through sound. Proceedings of ACM CHI ´03, (2003). Fort Lauderdale Florida, pp. 798-799
10. Sjostrom, C. Using haptics in computer interfaces for blind people. Proceeding of the ACM CHI ´01, Vol 3, 1, (2001), pp. 245-246
11. Siegfried, R. A scripting language to help the blind to program visually. ACM SIGPLAN Notices, Vol 37 (2), pp.53-56, February 2002
12. Smith, A.C., Francioni, J.M. & Matzek, S.D. A Java Programming tool for students with visual disabilities. Proceedings of ACM ASSETS ´00, pp. 142-148, 2000

eDiab: A System for Monitoring, Assisting and Educating People with Diabetes

L. Fernández-Luque[1], J.L. Sevillano[1,*], F.J. Hurtado-Núñez[2], F.J. Moriana-García[3], F. Díaz del Río[1], and D. Cascado[1]

[1] Robotics & Computer Technology for Rehabilitation Laboratory.-
University of Sevilla, Spain
[2] Clinical Psychologist - Specialist in Diabetic Education.
Spanish Society of Diabetes (SED) - Diabetes Education Study Group
[3] MSc Computer Science - eHealth Consultant
[*] Corresponding author: ETS Ingeniería Informática. University of Sevilla, Spain.
Avda. Reina Mercedes s/n. 41012, Sevilla
jlsevillano@us.es

Abstract. In this paper, a system developed for monitoring, assisting and educating people with diabetes, named eDiab, is described. A central node (PDA or mobile phone) is used at the patient's side for the transmission of medical information, health advices, alarms, reminders, etc. The software is adapted to blind users by using a screen reader called Mobile Speak Pocket/Phone. The glucose sensor is connected to the central node through wireless links (Zigbee/Bluetooth) and the communication between the central node and the server is established with a GPRS/GSM connection. Finally, a subsystem for health education (which sends medical information and advice like treatment reminder), still under development, is briefly described.

1 Introduction

The Diabetes Mellitus is one of the most common chronic diseases worldwide, characterised by a state of chronic hyperglycaemia (raised blood sugar). Diabetes Mellitus affects almost 6% of the population in Europe [1,2]. It is also one of the main causes of mortality in the world [3] and responsible for serious health complications [4,5,6,7] such as cardiac and renal diseases, depressions, blindness, etc. The Diabetes has a very deep impact in patient's life: glucose controls, insulin, dietetic restrictions, fears, etc. The management of the disease needs a lot of discipline and self-care by the patient, because only healthy habits could break down its progression and reduce the risk of these complications [8]. A good diabetic control is affected by lifestyle factors (diet, exercise, etc.) and self-care (treatment compliance, self-monitoring, etc.). The key component in empowering patients to manage their own diabetes is education.

The objective of a diabetes education program is to provide people with diabetes the knowledge, skills and tools to manage their disease and avoid the complications associated with the diabetes. Nowadays, educational programs include personal

K. Miesenberger et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 1342 – 1349, 2006.

meetings and lessons provided by diabetes educators (nurses, dieticians, doctors, psychologists, etc.).

Over the last years there has been an explosion of web pages with information for diabetic people. However, the inclusion of new technologies in educational programs is very low; a systematic review [9] of diabetes education has found that only 6% of the time in educational programs is spent in Computer-assisted instruction. Systematic reviews of computer-assisted education programs [10,11,12,13] have found evidences that emerging IT may improve diabetes care.

The use of new information and communication technologies in diabetes care has been focused on telemedicine projects. There are many telemedicine projects for diabetic people [14,15,16,17,18,19]. The aims of most of these projects were telemonitoring the patient's glucose [15,16,17,18] and tele-ophthalmology for the prevention of diabetic retinopathy, one of the main causes of blindness [19]. The first goal of this monitoring is the adjustment of the insulin treatment. One example of these projects is the M2DM project [16], in which a system has been developed for telemonitoring, alarms, decision support, electronic clinical record, SMS reminders, etc. Another interesting project is IDEATEL [15], focused on telemonitoring and education. In this project a complete system has been developed to control the patients. The system of IDEATEL includes: glucometers, blood pressure sensors, video conference, emails, educational resources via web, etc. The use of new technologies to improve the methods for delivering diabetes education is encouraged by the International Federation of Diabetes.

In our analysis of the state-of-art in telemedicine for diabetes, no works have been found that include blind people as users. This is a very important problem, since the diabetes Mellitus is the second cause of blindness in elderly people and the first one in young and adult people [5]. However, despite the high number of assistive products developed for the visually impaired [20,21], we have not found any work merging telemonitoring diabetes and assistance to visually impaired people. Unfortunately, this is a usual lack in telemedicine, although we think that there are no reasons to separate Assistive Technologies from Telemedicine. Furthermore, in telemedicine projects for diabetic people new technologies like mobile phones and PDAs are only used for the transmission of glucose measures or retina images and communication between doctors and patients. As far as we are concerned, these systems could also be used for the education of people suffering from diabetes. For instance, SMS and MMS could be used to improve the healthy habits of the patients by sending them personalized educational information.

In this paper, we propose a system that merges monitoring and educating people with diabetes, as well as assisting those users that are also visually impaired, probably due to diabetes complications.

2 Description

In the project eDiab an Information System for the Management of Diabetic People is being developed. This project has been designed with a multidisciplinary view: assistive technologies, telemedicine and diabetic education. The project has been divided into two subsystems: Subsystem for Monitoring and Controlling of Diabetes and Subsystem for Health Education.

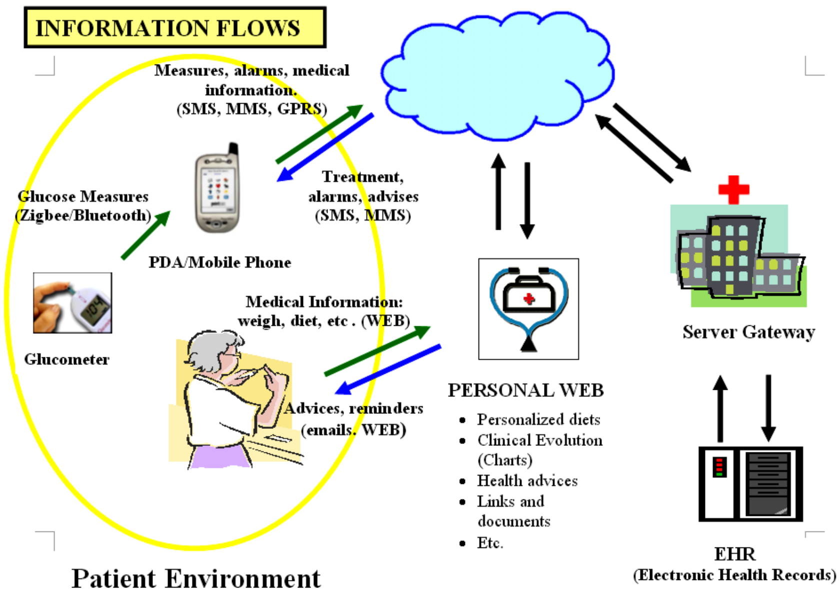


Fig. 1. Information flow of eDiab

2.1 Subsystem for Monitoring and Controlling

The subsystem for monitoring and controlling of Diabetes is the responsible for the transmission and storage of measures of glucose levels and other data from the patient such as information about diet, weight, etc. At the patient side there is a central node, PDA (for young users) or a mobile phone (for elder users). These devices are designed for the transmission of medical information, health advices, alarms, reminders, access to medical information, etc. The software of both devices is adapted to blind users by using an application called Mobile Speak Pocket/Phone [22], a screen reader solution for mobile phones that allows the access to most of the device functionality. The PDA has more features than a mobile phone, but its complexity is higher, so this device is mainly developed for young users. The design of the system for the PDA includes note reminders, a data base of glucose measures, generation of charts, etc.; these applications have been developed using .NET technology in a Pocket PC O.S. The PDA uses Zigbee/802.15.4 on the connection between the glucose sensor and the PDA, since this relatively new wireless sensor technology is the best option for battery saving and low cost [25]. This connection has been developed using Zigbee-RS232 modules of Telegesis Inc. [23], with the communication established by AT-commands. The first prototype had been developed with the Telegesis development board, but this development board was not a good choice because it is not battery powered and it needs to be powered with an adaptor. In order to overcome this drawback a new board with a small size and battery powered has been developed using Zigbee-RS232.

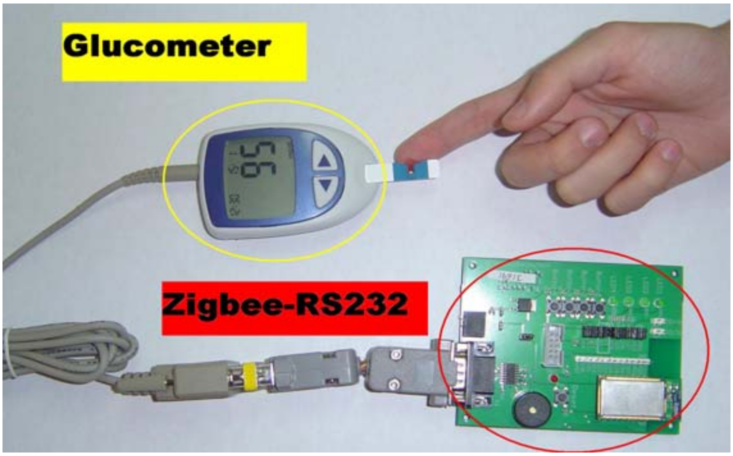


Fig. 2. Sensor connection with the Zigbee module

The communication between the patient and the server is established with a GPRS/GSM connection, using a Nokia 6630 connected to the PDA. Our current prototype is working in a HP iPAQ device; this PDA has a GPRS connection using a mobile phone connected via Bluetooth. However, the next prototypes will be developed in a Pocket PC with GPRS connection.

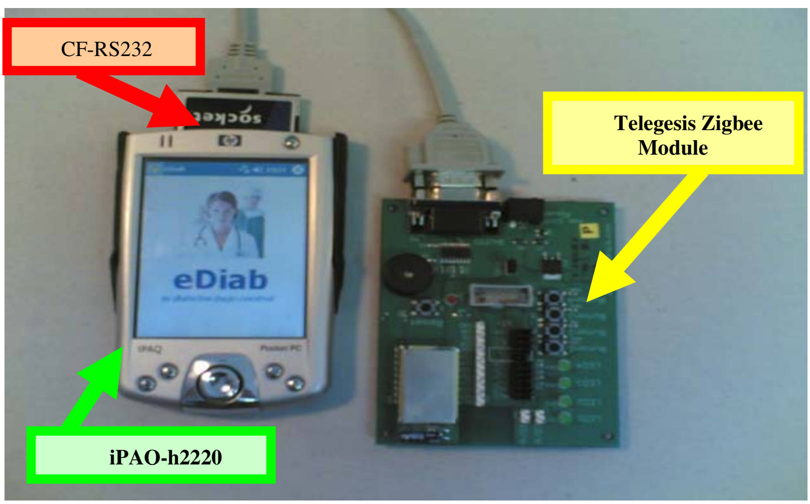


Fig. 3. PDA monitoring system

As described above, the mobile device used in the first prototype is a Nokia 6630 [24] with UMTS and Symbian O.S. The application software for this mobile has been developed using Java technology, J2ME. This device works as a gateway between the servers and the user. This bidirectional communication is based on GPRS/GSM, although the following prototypes could use UMTS. We have also used Bluetooth for the communication between the sensor and the mobile phone; in this case, the system uses a serial port-Bluetooth adapter.

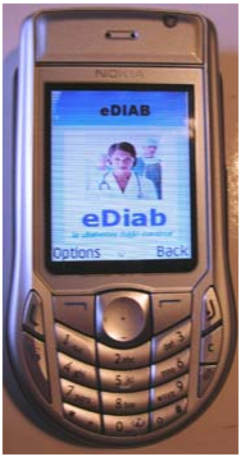


Fig. 4. eDIAB Application for Pocket PC

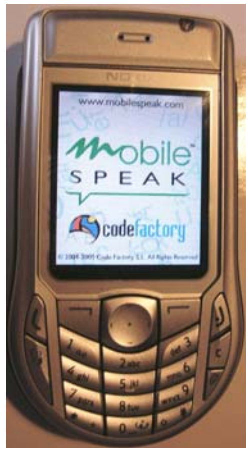


Fig. 5. Mobile Speak, Screen Reader Application

2.2 Subsystem for Health Education

The subsystem for Health Education has two components: one personal web for the patient (currently under development) and one system to send SMS and MMS with medical information and advice. The web allows the patient to access to the medical information (charts, analysis, etc.) and to send emails to the doctors and professionals. The web also shows advice for the patient, which are personalized and generated automatically. A smart system studies the information about the patient and provides customized information and tailored view of it in the personal web; for example, if the patient doesn't practise sports, the web will show information about the benefits of sports. This system also provides SMS/MMS features for health education. The server could send SMS/MMS to remind about the treatments and to give advice. This feature is being developed for improving the healthy habits of the patients. Two main target groups are considered: children, who are very receptive to this kind of system, and people with type II diabetes (the most common form of diabetes), who may need more advice in healthy living because most of them have been diagnosed with diabetes when they are adults.

3 Future Work and Challenges

Currently, the Subsystem for Monitoring and Controlling is being tested with end users in order to test the usability and to improve the system in the next prototype. The subsystem of Health Education has been designed and is currently being developed; we hope that it will be fully developed during the academic course 2006-2007. After this, the system will be tested on diabetic patients, including blind people, and diabetic educators.

There are several interesting areas for further research related to Information Technologies in Diabetic Care, especially in the educational area. For example, we have found projects [26] where animated agents were used to educate patients who had suffered brain stroke. We think that this approach has a high potential of applicability to the education of patients with chronic diseases. Another interesting area is the use of Artificial Intelligence techniques to personalize the presentation of documents [27,28] according to the patient's characteristics and interests.

Probably the more interesting challenge is the integration of eDiab into an Ambient Intelligent System. Ambient Intelligence has an enormous potential through the creation of "intelligent" environments able to proactively adapt to humans, as well as to serve their needs and goals. In this case, the autonomy and quality of life of diabetic people would be enhanced through a proactive education, monitoring and self-care system providing an integral care of this very important chronic disease. The use of wireless links guarantees ubiquitous access to medical data (e.g. sensors) and advice and/or medical information. Also, the ubiquitous access to electronic clinical records and the application of smart techniques would allow personalizing and adapting advice and medical information to the user's characteristics (age, abilities, etc.). Future work will address the improvement of the system described in this paper in order to allow *interoperability* with other devices. For instance, a person using eDiab and entering a hospital or school may be interested in interchanging services and/or information. We are currently considering the use of one of the available architectures (like UPnP or Jini) for dynamic service discovering, service description and service control.

4 Conclusions

In this paper, we describe a system named eDiab, developed for monitoring, assisting and educating people with diabetes. The main contribution of this system is that it is (to the best of our knowledge) the first one in merging monitoring and education of diabetic people with assistance to visually impaired people. This feature is especially useful considering that the Diabetes Mellitus is the second cause of blindness in the elderly and the first one in young and adult people. Although the system is still under development, several interesting characteristics can already be presented.

Acknowledgements

This work has been partially supported by projects TIN2005-08976 and TIC2003-08164-C03-02.

References

1. King H, Aubert RE, Herman WH. "Global Burden of Diabetes 1995-2025: Prevalence, Numerical Estimates and Projections" – World Health Organization, Diabetes Care 2 1 :1 4 1 4–1431, 1998
2. Gan D, et al. "Diabetes Atlas". International Diabetes Federation, 2003 `http://www.eatlas.idf.org/`

3. Gojka Roglic, et al. “The Burden of Mortality Attributable to Diabetes” World Health Organization, Diabetes Care 28:2130–2135, 2005
4. Rosenzweig JL, Weinger K, et al. “Use of a disease severity index for evaluation of healthcare costs and management of comorbidities of patients with diabetes mellitus”. 2002, Office for Disease Management, Joslin Diabetes Center, Boston
5. Rahmani B, Tielsch JM, et al. “The cause-specific prevalence of visual impairment in an urban population”. The Baltimore Eye Survey. Ophthalmology 1996; 03:1721-1726
6. Haffner SM, et al. “Mortality from coronary heart disease in subjects with type 2 diabetes and in nondiabetic subjects with and without prior myocardial infarction”. New England Journal of Medicine 1998; 339:229-234
7. Lin EH, Katon W, et al. “Relationship of depression and diabetes self-care, medication adherence, and preventive care”. Diabetes Care 2004, 27:2154-60
8. Norris SL, Engelgau MM, et al. Effectiveness of self-management training in type 2 diabetes. A systematic review of randomized controlled trials. Diabetes Care 2001; 24: 561-87.
9. S. L. Norris, J. Lau, et al. Self-Management Education for Adults with Type 2 Diabetes: A meta-analysis of the effect on glycemic control. Diabetes Care Journal, Number 25, 2002, pp 1159-1171
10. Montani S, Bellazzi R, Quaglini S, D’Annunzio G., et al. Meta-analysis of the effect of the use of computerbased systems on the metabolic control of patients with diabetes mellitus. Diabetes Technol Ther 2001; pp 347–356
11. Jackson, Chandra L.; Bolen, et al. A Systematic Review of Interactive Computer-assisted Technology in Diabetes Care. Interactive Information Technology in Diabetes Care. Journal of General Internal Medicine, Volume 21, Number 2, February 2006, pp 105-110
12. D. Lewis. Computer-based Approaches to Patient Education A Review of the Literature. Journal of American Medical Informatics Association. 1999 Jul–Aug; pp 272–282
13. G.T. McMahon, H.E. Gomes, et al. Web-Based Care Management in Patients with Poorly Controlled Diabetes. Diabetes Care Journal; 28:1624-1629, 2005
14. Klonoff DC. Diabetes and telemedicine. Is the technology sound, effective, cost-effective and practical? Diabetes Care 2003; 26: 1626-28
15. Justin Starren, et al. Columbia University’s Informatics for Diabetes Education and Telemedicine (IDEATel) Project. Journal of American Medical Informatics Association. 2002; 9:25–36
16. Bellazi R., Arcelloni M., et al. Design, methods, and evaluation directions of a multi-access service for the management of diabetes mellitus patients. Diabetes Technology & Therapeutics, 2003, Vol. 5, No. 4: 621-629
17. European Project Inca, “Intelligent Control Assistant for Diabetes”, 5th EC Framework Programme, `http://www.ist-inca.org`
18. Joe Cafazzo, “Tele-management of Diabetic Hypertension: The use of the digital medical diary”, eHealth Conference 2005
19. Williamson T.H., et al., “Telemedicine and computers in diabetic retinopathy screening”, British Journal of Ophthalmology. 1998;82;5-6
20. J. Tejedor, D. Bolaños, et al., Multimedia Medicine Consultant for Visually Impaired People. LNCS 3118 (ICCHP 2004), 564-570
21. A.H.M.N. Khan, S.N. Chowdhury “Computer for Persons with Visually Impairment: A Door to Education, Information and Employment”. LNCS 3118 (ICCHP 2004), 571-575
22. Mobile Speak Software. `http://www.codefactory.es/`
23. Telegesis Inc, `http://www.telegesis.com`
24. Nokia 6630 Technical Characteristics, `http://www.nokia.co.uk/nokia/0,,59880,00.html`

25. J.A. Gutierrez, E.H. Callaway, R. Barrett. Low-Rate Wireless Personal Area Networks: Enabling Wireless Sensor Networks with IEEE 802.15.4. IEEE Press, 2004
26. Holden, M.K., Dyar, T., Schwamm, L., et al. (2003). Home-based Telerehabilitation using a virtual environment system. In Burdea GC, Thalmann D, & Lewis JA (Eds.): Proceedings of the 2nd International Workshop on Virtual Rehabilitation, 2003, pp 4-12
27. Hirst G, DiMarco C, Hovy E, Parsons K. Authoring and generating health-education documents that are tailored to the needs of the individual patient. Proceedings of the Sixth International Conference on User Modeling, Sardinia, Italy, 1997, 107-118
28. DiMarco, C., P. Bray, D. Covvey, D. Cowan, V. DiCiccio, J. Lipa, and E.H. Hovy. 2005. Authoring and Generation of Tailored Preoperative Patient Education Materials. Proceedings of the Workshop on Personalisation for e-Health at the User Modeling (UM) Conference. Edinburgh, UK

Author Index

Abascal, Julio 120, 337, 777
Abou-Zahra, Shadi 192
Adjouadi, Malek 926, 1094, 1163
Agelfors, Eva 579
Agnoletto, Andrea 524
Aguayo, Fernando 1334
Aguilera, Santiago 675
Aitken, Jonathan 24
Akita, Jun-ichi 1251
Albert, Peter 1118
Alonso, Fernando 1031, 1071, 1207
Alonso Jr., Miguel 905, 1163
Altmanninger, Kerstin 128
Al-Wabil, Areej 817
Andjomshoaa, Amin 1147
Andrich, Renzo 492, 524
Antona, Margherita 421
Arato, Andras 1055
Archambault, Dominique 369, 372, 1191
Arch-Tirado, Emilio 691
Arenas, Jerónimo 675
Arrue, Myriam 120
Artigas, José I. 744, 777
Aslaksen, Finn 258
Atkinson, Matthew T. 380
Aupetit, Sébastien 500
Ayala, Melvin 926

Baloian, Nelson 1318, 1326
Barajas-Montiel, Sandra E. 691
Barbier, Ander 279
Bargi Rangin, Hadi 230
Barreto, Armando 905, 1163
Bártek, Luděk 114
Bartsch, Martin 488
Batusic, Mario 1235
Batušić, Mario 1306
Beck, Christian 886
Becker, Josef 322
Benavídez, Carlos 199, 341
Berglund, Tomas 701
Bernareggi, Cristian 1223
Beskow, Jonas 579
Bickley, Corine 643
Billi, M. 407
Birnbaum, Martha 643
Blenkhorn, Paul 855, 874
Bognar, Gabor 1055
Boll, Susanne 69
Bothe, Hans-Heinrich 575
Boutsakis, Evagelos 421
Bowick, Lisa 874
Boyer, John J. 1215
Brajnik, Giorgio 156
Broekman, Christiane 587
Bruhn, Lars 532
Brzoza, Piotr 1087
Buday, Laszlo 1055
Bühler, Christian 184, 322
Buiza, C. 759
Bulatov, Vladimir 1243
Burger, Brianne 643
Burger, Dominique 85, 242, 245
Burzagli, L. 407

Cantón, Paloma 536
Caracciolo, Antonio 492
Casas, Oscar 1006
Casas, Roberto 724, 744, 777, 1006
Cascado, D. 1342
Challis, Ben P. 1110
Chang, Tzyh-Chyang 516, 770
Chen, Ching-Horng 437
Chen, Li-Yuan 839
Chen, Ming-Chung 450, 839, 919, 966, 981
Chin, Craig 905
Christensen, Lars Ballieu 1102
Chu, Chi-Nung 450
Chung, Yun Koo 469
Cko, Chien-Chuan 839
Coughlan, James 1294
Crombie, David 1, 16
Cunningham, Stuart 882

Da Dalt, Roberto 524
Daunys, Gintautas 12, 962

Dautenhahn, Kerstin 785
Davis, Megan 785
Del Zanna, Giovanni 492
Desbuquois, Catherine 85
Di Rienzo, Marco 492
Díaz del Río, F. 1342
Domingo, Muriel Garreta 1313
Draffan, E.A. 855

Eberhardt, Gergely 1055
Edelmayer, Georg 886, 1021
Ehrenstrasser, Lisa 415
Emiliani, P.L. 407
Encelle, Benoit 1223
Enderby, Pam 882
Engelen, Jan 1063
Etxeberria, I. 759
Evans, David Gareth 855, 874
Evreinov, Grigori 353, 1029, 1079
Evreinova, Tatiana G. 353

Falcó, Jorge L. 724, 744, 777, 831
Falcó, José M. 724
Fardal, Frank 258
Fejt, Jan 954
Fejtová, Marcela 954
Fels, Deborah I. 24
Fernández-Luque, L. 1342
Ferraro, F. 1171
Ferres, Leo 1122
Fischer, Patrick 287
Fitzpatrick, D. 1199
Flyen, Are 396
Fraiture, Martine 587
Freda, C. 1171
Fuertes, José L. 199, 341, 599, 1031, 1071, 1207
Fugger, Erwin 699
Fujimoto, Madoka 943
Fujiyoshi, Akio 801
Fujiyoshi, Mamoru 801
Fukuda, Ryoji 1139

Gaal, Andrea 1306
Gabbanini, F. 407
Galdona, N. 759
Gale, Alastair 480
Gardner, John A. 1215, 1243
Gaucher, Pierre 500
Geiszt, Zoltán 823
Georgalis, Yannis 388
Geramani, Konstantina N. 16
Gil Rodríguez, Eva Patrícia 1313
Giust, Renzo 524
Gjøsæter, Terje 40
Gondo, Yasuyuki 752
González, Ángel Lucas 536, 599, 1031, 1207
González González, Manuel 847
González, Marcos 223
González, Martin 223
González, M.F. 759
Goodwin Olsen, Morten 234
Gotoh, Hiroshi 1265
Gower, Valerio 492
Gracia, Héctor 744
Grammenos, Dimitris 388
Gravenhorst, Claus 92
Graziani, P. 407
Greene, Cara Nicole 793
Green, Phil 882
Gucukoglu, Sabahattin 380
Guenaga, Mari Luz 279
Guillon, Benoît 85
Guillou, Pierre 245
Guo, Jun-Nan 839
Gut, Margarethe 1223
Gutiérrez, Emmanuelle 199, 341

Haffner, Christoph 349
Haindl, Philipp 1235
Häkkinen, Markku T. 96
Hanakovič, Tomáš 1231
Hara, Daisuke 595
Hartrumpf, Sven 207
Hawley, Mark S. 882
Heck, Helmut 184, 322
Helbig, Hermann 207
Hengstberger, Barbara 361
Herramhof, Sandor 164, 172, 215
Herrera, Andres 926
Heuten, Wilko 69
Hierrezuelo Sabatela, Lisset 894
Hill, Magí Almirall 1313
Hofmann-Wellenhof, Rainer 732
Holmesland Snaprud, Mikael 234
Holzinger, Andreas 732, 989, 999
Homann, Jürgen 532
Hori, Junichi 950
Horiuchi, Yasuo 863

Horstmann, Mirko 69
Huang, Juei-Fen 516, 770
Hurtado-Núñez, F.J. 1342
Hyyppä, Kalevi 701

Ichikawa, Akira 863, 902
Ienaga, Takafumi 1265
Ikeda, Yumi 943
Inagaki, Hiroki 752
Inoue, Kaoru 943
Ioannidis, George 16
Ishizaki, Shun 53
Ito, Kiyohide 1251, 1273
Ito, Yuko 943
Itoh, Kazuyuki 902, 973
Izumi, Chie 667

Jacko, Julie A. 1163
James, A. 855
Jardí, Llorenç Sabaté 1313
Jarochowski, Bart 473
Jašková L'udmila 552
Jenge, Constantin 207
Jessel, Nadine 1223
Johnson, Marianne 874
Juhasz, Zoltan 1055

Kahana, Boaz 766
Kahana, Eva 766
Kanahori, Toshihiro 48
Kanda, Kazuyuki 595
Karim, Shuaib 77, 1147
Karlsson, Inger 579
Károlyi, Péter 823
Karshmer, Arthur 635
Karshmer, Arthur I. 1175
Kato, Norihiko 1131
Kawahigashi, Kana 623
Kawamura, Syouhei 607
Kawane, Fukashi 1183
Kawano, Sumihiro 667
Kehagias, Dionisis 306
Kerscher, George 100
Kewley, Jo 579
Kim, Daeyoung 716
Kim, HyungJun 473, 709, 1014
Kim, Kyung-il 469
Kim, Yung Bok 263, 716
Kimura, Tsutomu 595
Kimura, Youko 1265

Kitajima, Muneo 136
Klaus, Joachim 1306
Kobayashi, Makoto 870
Koch, Johannes 172
Kocsor, András 615
Komada, Toshihiko 1183
Komatsu, Takanori 1251
Komeda, Takashi 935, 943
Kopeček, Ivan 61, 114
Koskinen, Simo 701
Kurokawa, Takao 607, 667

Lamy, Cécile 587
Lauruska, Vidas 12
Lawrence, Adrian E. 380
Lee, Jiann-Der 516, 770
Leitner, Marie-Luise 271
Leopold, Matthias 6
Leporini, Barbara 176
Lin, Yun-Lung 981
Lindgaard, Gitte 1122
Liu, Ji-Fu 839
Loeding, Barbara 635
Lórencz, Róbert 1047
Lovegreen, Loren 766
Lu, Chiu-Ping 437
Luo, An-Jun 839
Lurf, Robert 415

Machin, Colin H.C. 380
Macías, Javier 675
Macías Guarasa, Javier 894
Magyar, Viktor 508
Mahr, Michael 992, 999
Mamada, Kazuhiko 1273
Manduchi, Roberto 1294
Mandujano, Mario 691
Marco, Álvaro 724, 744, 777
Mariscal, Gonzalo 536, 599
Markus, Norbert 1055
Martínez Calvo, Francisco 16
Martín, Jose L. 675
Martínez, Loïc 199, 341, 1031, 1071, 1207
Martín-Sánchez, José L. 894
Massaro, Dominic W. 809
Masuda, Fumihiko 429
Matausch, Kerstin 361
Matsui, Hirokazu 1131
Matsumoto, Michito 1265

Mayer, Peter 1021
Mayrhofer, Thomas 488
McMullin, Barry 106
Meng, Ling-Fu 437
Miene, Andrea 69
Miesenberger, Klaus 32, 253, 271, 329, 361, 403, 488, 560, 1235
Mina, Stefan 415
Minagawa, Hiroki 627
Miura, Jun 623
Miyagi, Manabi 863
Miyakawa, Masahiro 1259, 1287
Miyakawa, Michio 950
Moço, Victor 1191
Mogor, Emil 1055
Monmarché, Nicolas 372, 500
Morandell, Martin 699
Moriana-García, F.J. 1342
Morimoto, Kazunari 607, 667
Mourouzis, Alexandros 421
Mühlehner, Mario 442
Murai, Yasuyuki 1259, 1287
Muro, Carmen 831

N.S. Andreasen Struijk, Lotte 913
Na, Jongwhoa 1298
Nagai, Nobuyuki 1259, 1287
Nagashima, Yuji 659
Nagy, Marek 1231
Nagy, Zoltan 1055
Nakamura, Miyoko 53
Nakazono, Kaoru 659
Namatame, Miki 136
Namgoong, Hyun 469
Nehaniv, Chrystopher 785
Németh, Géza 651
Nietzio, Annika 184
Nischelwitzer, Alexander K. 989, 992, 999
Nishi, Atsushi 1139
Nishida, Masafumi 863
Nishioka, Tomoyuki 136
Nitta, Osamu 935
Nomura, Yoshihiko 1131
Nussbaum, Gerhard 457, 488
Nuttin, Marnix 1021
Nytun, Jan Pettersen 40

Ogawa, Madoka 752
O'Grady, Mary 1306
Ohnishi, Noboru 627
Okamoto, Akira 902
Okamoto, Makoto 1251
Okikawa, Shigeru 935
Oliver, Javier 279
Ono, Tetsuo 1251
Ortner, Daniela 253, 271
Ossmann, Roland 403
Osswald, Rainer 207
Ošlejšek, Radek 61

Paczolay, Dénes 615
Pagliara, S. 1171
Palazuelos-Cagigas, Sira E. 894
Palazuelos, Sira E. 675
Palmer, Rebecca 882
Panek, Paul 1021
Paredes Garrido, María Dolores 847
Parush, Avi 1122
Paternò, Fabio 176
Pepino, A. 1171
Perlick, Olaf 184
Petrie, Helen 152, 215
Petz, Andrea 560
Piekkari, Jouni 701
Pillai, Anand Balachandran 234
Pino Morillas, Oscar 847
Pintado, Ignacio 223
Plaza, Inmaculada 831
Ponsard, Christophe 314, 587
Powell, Stuart D. 785
Power, Christopher 1155
Prazak, Barbara 415, 699
Pressl, Bettina 1280
Prinz, Andreas 40
Pühretmair, Franz 295, 329
Purdy, Kevin 480
Puret, Arnaud 500

Quilez, Marcos 1006

Raike, Antti 568
Raisamo, Roope 353
Ramanauskas, Nerijus 962
Rashid, Raisa 24
Renzelberg, Gerlinde 532
Reyes-García, Carlos A. 691
Ribera, Jordi Planella 1313
Rissanen, Jyrki 740
Rivera, Cristóbal 223

Roberts, Shelley 1122
Rodríguez Fórtiz, María José 847
Rodríguez Parra, Encarnación 847
Romero, Borja 1006
Rönnbäck, Sven 701
Roy, Armando 831
Ruemer, Reinhard 32
Ruiz, Carlos 536, 599
Ruiz, Eddy 1094
Ryu, DaeHyun 473, 709, 1014

Saitoh, Yoshiaki 950
Sakano, Koji 950
Sakashita, Yudai 627
Salvi, Giampiero 579
Sammer, Peter 732
Sánchez, Jaime 1318, 1326, 1334
Sarkar, Sudeep 635
Savidis, Anthony 388
Savitch, Nada 144
Schrepp, Martin 287
Schweikhardt, Waltraud 1223
Scorcia, Antonio 176
Seçkin, Gul 766
Seisenbacher, Gottfried 886
Sepchat, Alexis 372
Sevillano, J.L. 1342
Shen, Huiying 1294
Shi, Fangmin 480
Shibata, Masahiro 1265
Shih, Chun-Hua 770
Shimada, Nobutaka 623
Shimada, Shigenobu 1039
Shimizu, Yutaka 1039
Shimojo, Makoto 1039
Shinohara, Masami 1039
Shiose, Takayuki 1273
Shirai, Yoshiaki 623
Sik-Lányi, Cecília 508, 823
Sischka, Dirk 322
Skjelten Tveit, Merete 40
Slimane, Mohamed 372, 500
Snaprud, Mikael 40, 258
Snoeck, Vincent 314
Sommer, Sascha M. 306
Spinczyk, Dominik 1087
Sproger, Bernd 999
Stanley, Paul B. 1175
Štěpánková, Olga 954
Stephanidis, Constantine 388, 421
Stoeger, Bernhard 1235
Strauss, Christine 271
Strobbe, Christophe 164, 172, 215, 544
Strobl-Mairhofer, Christoph 488
Sugiura, Tokuhiro 1131
Suzuki, Keisuke 943
Suzuki, Masakazu 48, 1183
Suzuki, Tadashi 943
Szabo, Héctor 1071

Takahashi, Yoshiyuki 935, 943
Takeuchi, Yoshinori 627
Tanimura, Atsuko 943
Tatsumi, Hisayuki 1259, 1287
Terauchi, Mina 659
Thomas, Neil 579
Tjoa, A. Min 77, 1147
Tollefsen, Morten 396
Torihara, Shinichi 53
Tóth, Bálint 651
Toyoda, Nobuyuki 1265
Tsuchijima, Masahiro 943
Tzeng, Wen-Jeng 981
Tzovaras, Dimitrios 306

Ueda, Nami 53
Ulltveit-Moe, Nils 184, 234
Ungier, Leon 1215
Urdaneta, E. 759

Valjakka, Sari 306
Van Hees, Kris 1063
Vaspori, Terez 1055
Velasco, Carlos A. 164, 172, 215
Vidau, Agueda 223
Vigo, Markel 120
Vlachogiannis, Evangelos 164, 172, 215

Wada, Chikamune 429
Wada, Takuya 53
Wakatsuki, Daisuke 627
Wald, Mike 683
Walshe, Esmond 106
Wang, Lu 1094
Weber, Gerhard 152, 215, 337, 349
Weimann, Kurt 215
Wieser, Manfred 1280
Wiethoff, Marion 306
Wilson, Stephanie 817
Winkler, Michael 298

Wöß, Wolfram 128, 298
Wu, Shwu-Jiuan 516, 770
Wu, Ting-Fang 450, 919, 966

Yagi, Yuki 1131
Yamaguchi, Katsuhito 1183
Yang, Ming-Jen 770
Yang, Ruiduo 635
Yanguas, J.J. 759
Yasukouchi, Takehito 1265
Yasumura, Michiaki 461, 902
Yeh, Chih-Ching 450, 981
Yeh, Yao-Ming 981
Yfantidis, Georgios 1079
Yoshida, Megumi 461
Yoshida, Ryo 461

Zagler, Wolfgang 886, 1021
Zahradnický, Tomáš 1047
Zanfardino, F. 1171
Zaphiris, Panayiotis 144, 817