LNCS 6771

Michael J. Smith
Gavriel Salvendy (Eds.)

Human Interface and the Management of Information

Interacting with Information

Symposium on Human Interface 2011
Held as Part of HCI International 2011
Orlando, FL, USA, July 2011, Proceedings, Part I

1 Part I

HCI2011 INTERNATIONAL

Springer

Lecture Notes in Computer Science 6771

Michael J. Smith Gavriel Salvendy (Eds.)

Human Interface and the Management of Information

Interacting with Information

Symposium on Human Interface 2011
Held as Part of HCI International 2011
Orlando, FL, USA, July 9-14, 2011
Proceedings, Part I

Springer

Volume Editors

Michael J. Smith
University of Wisconsin-Madison
Department of Industrial and Systems Engineering
Center for Quality and Productivity Improvement
Madison, WI, USA
E-mail: mjsmith@cae.wisc.edu

Gavriel Salvendy
Purdue University
School of Industrial Engineering
West Lafayette, IN, USA
and
Tsinghua University
Department of Industrial Engineering
Beijing, P.R. China
E-mail: salvendy@purdue.edu

ISSN 0302-9743 e-ISSN 1611-3349
ISBN 978-3-642-21792-0 e-ISBN 978-3-642-21793-7
DOI 10.1007/978-3-642-21793-7
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2009928850

CR Subject Classification (1998): H.5, K.6, H.3-4, C.2, H.4.2, J.1

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

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Foreword

The 14th International Conference on Human–Computer Interaction, HCI International 2011, was held in Orlando, Florida, USA, July 9–14, 2011, jointly with the Symposium on Human Interface (Japan) 2011, the 9th International Conference on Engineering Psychology and Cognitive Ergonomics, the 6th International Conference on Universal Access in Human–Computer Interaction, the 4th International Conference on Virtual and Mixed Reality, the 4th International Conference on Internationalization, Design and Global Development, the 4th International Conference on Online Communities and Social Computing, the 6th International Conference on Augmented Cognition, the Third International Conference on Digital Human Modeling, the Second International Conference on Human-Centered Design, and the First International Conference on Design, User Experience, and Usability.

A total of 4,039 individuals from academia, research institutes, industry and governmental agencies from 67 countries submitted contributions, and 1,318 papers that were judged to be of high scientific quality were included in the program. These papers address the latest research and development efforts and highlight the human aspects of design and use of computing systems. The papers accepted for presentation thoroughly cover the entire field of human–computer interaction, addressing major advances in knowledge and effective use of computers in a variety of application areas.

This volume, edited by Michael J. Smith and Gavriel Salvendy, contains papers in the thematic area of human interface and the management of information (HIMI), addressing the following major topics:

- Design and development methods and tools
- Information and user interfaces design
- Visualisation techniques and applications
- Security and privacy
- Touch and gesture interfaces
- Adaptation and personalisation
- Measuring and recognising human behaviour

The remaining volumes of the HCI International 2011 Proceedings are:

- Volume 1, LNCS 6761, Human–Computer Interaction—Design and Development Approaches (Part I), edited by Julie A. Jacko
- Volume 2, LNCS 6762, Human–Computer Interaction—Interaction Techniques and Environments (Part II), edited by Julie A. Jacko
- Volume 3, LNCS 6763, Human–Computer Interaction—Towards Mobile and Intelligent Interaction Environments (Part III), edited by Julie A. Jacko
- Volume 4, LNCS 6764, Human–Computer Interaction—Users and Applications (Part IV), edited by Julie A. Jacko

- Volume 5, LNCS 6765, Universal Access in Human–Computer Interaction—Design for All and eInclusion (Part I), edited by Constantine Stephanidis
- Volume 6, LNCS 6766, Universal Access in Human–Computer Interaction—Users Diversity (Part II), edited by Constantine Stephanidis
- Volume 7, LNCS 6767, Universal Access in Human–Computer Interaction—Context Diversity (Part III), edited by Constantine Stephanidis
- Volume 8, LNCS 6768, Universal Access in Human–Computer Interaction—Applications and Services (Part IV), edited by Constantine Stephanidis
- Volume 9, LNCS 6769, Design, User Experience, and Usability—Theory, Methods, Tools and Practice (Part I), edited by Aaron Marcus
- Volume 10, LNCS 6770, Design, User Experience, and Usability—Understanding the User Experience (Part II), edited by Aaron Marcus
- Volume 12, LNCS 6772, Human Interface and the Management of Information—Interacting with Information (Part II), edited by Gavriel Salvendy and Michael J. Smith
- Volume 13, LNCS 6773, Virtual and Mixed Reality—New Trends (Part I), edited by Randall Shumaker
- Volume 14, LNCS 6774, Virtual and Mixed Reality—Systems and Applications (Part II), edited by Randall Shumaker
- Volume 15, LNCS 6775, Internationalization, Design and Global Development, edited by P.L. Patrick Rau
- Volume 16, LNCS 6776, Human-Centered Design, edited by Masaaki Kurosu
- Volume 17, LNCS 6777, Digital Human Modeling, edited by Vincent G. Duffy
- Volume 18, LNCS 6778, Online Communities and Social Computing, edited by A. Ant Ozok and Panayiotis Zaphiris
- Volume 19, LNCS 6779, Ergonomics and Health Aspects of Work with Computers, edited by Michelle M. Robertson
- Volume 20, LNAI 6780, Foundations of Augmented Cognition: Directing the Future of Adaptive Systems, edited by Dylan D. Schmorrow and Cali M. Fidopiastis
- Volume 21, LNAI 6781, Engineering Psychology and Cognitive Ergonomics, edited by Don Harris
- Volume 22, CCIS 173, HCI International 2011 Posters Proceedings (Part I), edited by Constantine Stephanidis
- Volume 23, CCIS 174, HCI International 2011 Posters Proceedings (Part II), edited by Constantine Stephanidis

I would like to thank the Program Chairs and the members of the Program Boards of all Thematic Areas, listed herein, for their contribution to the highest scientific quality and the overall success of the HCI International 2011 Conference.

In addition to the members of the Program Boards, I also wish to thank the following volunteer external reviewers: Roman Vilimek from Germany, Ramalingam Ponnusamy from India, Si Jung “Jun” Kim from the USA, and Ilia Adami, Iosif Klironomos, Vassilis Kouroumalis, George Margetis, and Stavroula Ntoa from Greece.

This conference would not have been possible without the continuous support and advice of the Conference Scientific Advisor, Gavriel Salvendy, as well as the dedicated work and outstanding efforts of the Communications and Exhibition Chair and Editor of HCI International News, Abbas Moallem.

I would also like to thank for their contribution toward the organization of the HCI International 2011 Conference the members of the Human–Computer Interaction Laboratory of ICS-FORTH, and in particular Margherita Antona, George Paparoulis, Maria Pitsoulaki, Stavroula Ntoa, Maria Bouhli and George Kapnas.

July 2011 Constantine Stephanidis

Organization

Ergonomics and Health Aspects of Work with Computers

Program Chair: Michelle M. Robertson

Arne Aarås, Norway
Pascale Carayon, USA
Jason Devereux, UK
Wolfgang Friesdorf, Germany
Martin Helander, Singapore
Ed Israelski, USA
Ben-Tzion Karsh, USA
Waldemar Karwowski, USA
Peter Kern, Germany
Danuta Koradecka, Poland
Nancy Larson, USA
Kari Lindström, Finland
Brenda Lobb, New Zealand
Holger Luczak, Germany
William S. Marras, USA
Aura C. Matias, Philippines
Matthias Rötting, Germany
Michelle L. Rogers, USA
Dominique L. Scapin, France
Lawrence M. Schleifer, USA
Michael J. Smith, USA
Naomi Swanson, USA
Peter Vink, The Netherlands
John Wilson, UK

Human Interface and the Management of Information

Program Chair: Michael J. Smith

Hans-Jörg Bullinger, Germany
Alan Chan, Hong Kong
Shin'ichi Fukuzumi, Japan
Jon R. Gunderson, USA
Michitaka Hirose, Japan
Jhilmil Jain, USA
Yasufumi Kume, Japan
Mark Lehto, USA
Hirohiko Mori, Japan
Fiona Fui-Hoon Nah, USA
Shogo Nishida, Japan
Robert Proctor, USA
Youngho Rhee, Korea
Anxo Cereijo Roibás, UK
Katsunori Shimohara, Japan
Dieter Spath, Germany
Tsutomu Tabe, Japan
Alvaro D. Taveira, USA
Kim-Phuong L. Vu, USA
Tomio Watanabe, Japan
Sakae Yamamoto, Japan
Hidekazu Yoshikawa, Japan
Li Zheng, P. R. China

Human–Computer Interaction

Program Chair: Julie A. Jacko

Sebastiano Bagnara, Italy
Sherry Y. Chen, UK
Marvin J. Dainoff, USA
Jianming Dong, USA
John Eklund, Australia
Xiaowen Fang, USA
Ayse Gurses, USA
Vicki L. Hanson, UK
Sheue-Ling Hwang, Taiwan
Wonil Hwang, Korea
Yong Gu Ji, Korea
Steven A. Landry, USA
Gitte Lindgaard, Canada
Chen Ling, USA
Yan Liu, USA
Chang S. Nam, USA
Celestine A. Ntuen, USA
Philippe Palanque, France
P.L. Patrick Rau, P.R. China
Ling Rothrock, USA
Guangfeng Song, USA
Steffen Staab, Germany
Wan Chul Yoon, Korea
Wenli Zhu, P.R. China

Engineering Psychology and Cognitive Ergonomics

Program Chair: Don Harris

Guy A. Boy, USA
Pietro Carlo Cacciabue, Italy
John Huddlestone, UK
Kenji Itoh, Japan
Hung-Sying Jing, Taiwan
Wen-Chin Li, Taiwan
James T. Luxhøj, USA
Nicolas Marmaras, Greece
Sundaram Narayanan, USA
Mark A. Neerincx, The Netherlands
Jan M. Noyes, UK
Kjell Ohlsson, Sweden
Axel Schulte, Germany
Sarah C. Sharples, UK
Neville A. Stanton, UK
Xianghong Sun, P.R. China
Andrew Thatcher, South Africa
Matthew J.W. Thomas, Australia
Mark Young, UK
Rolf Zon, The Netherlands

Universal Access in Human–Computer Interaction

Program Chair: Constantine Stephanidis

Julio Abascal, Spain
Ray Adams, UK
Elisabeth André, Germany
Margherita Antona, Greece
Chieko Asakawa, Japan
Christian Bühler, Germany
Jerzy Charytonowicz, Poland
Pier Luigi Emiliani, Italy
Michael Fairhurst, UK
Dimitris Grammenos, Greece
Andreas Holzinger, Austria
Simeon Keates, Denmark
Georgios Kouroupetroglou, Greece
Sri Kurniawan, USA
Patrick M. Langdon, UK
Seongil Lee, Korea

Zhengjie Liu, P.R. China
Klaus Miesenberger, Austria
Helen Petrie, UK
Michael Pieper, Germany
Anthony Savidis, Greece
Andrew Sears, USA
Christian Stary, Austria
Hirotada Ueda, Japan
Jean Vanderdonckt, Belgium
Gregg C. Vanderheiden, USA
Gerhard Weber, Germany
Harald Weber, Germany
Panayiotis Zaphiris, Cyprus

Virtual and Mixed Reality

Program Chair: Randall Shumaker

Pat Banerjee, USA
Mark Billinghurst, New Zealand
Charles E. Hughes, USA
Simon Julier, UK
David Kaber, USA
Hirokazu Kato, Japan
Robert S. Kennedy, USA
Young J. Kim, Korea
Ben Lawson, USA
Gordon McK Mair, UK
David Pratt, UK
Albert “Skip” Rizzo, USA
Lawrence Rosenblum, USA
Jose San Martin, Spain
Dieter Schmalstieg, Austria
Dylan Schmorrow, USA
Kay Stanney, USA
Janet Weisenford, USA
Mark Wiederhold, USA

Internationalization, Design and Global Development

Program Chair: P.L. Patrick Rau

Michael L. Best, USA
Alan Chan, Hong Kong
Lin-Lin Chen, Taiwan
Andy M. Dearden, UK
Susan M. Dray, USA
Henry Been-Lirn Duh, Singapore
Vanessa Evers, The Netherlands
Paul Fu, USA
Emilie Gould, USA
Sung H. Han, Korea
Veikko Ikonen, Finland
Toshikazu Kato, Japan
Esin Kiris, USA
Apala Lahiri Chavan, India
James R. Lewis, USA
James J.W. Lin, USA
Rungtai Lin, Taiwan
Zhengjie Liu, P.R. China
Aaron Marcus, USA
Allen E. Milewski, USA
Katsuhiko Ogawa, Japan
Oguzhan Ozcan, Turkey
Girish Prabhu, India
Kerstin Röse, Germany
Supriya Singh, Australia
Alvin W. Yeo, Malaysia
Hsiu-Ping Yueh, Taiwan

Online Communities and Social Computing

Program Chairs: A. Ant Ozok, Panayiotis Zaphiris

Chadia N. Abras, USA
Chee Siang Ang, UK
Peter Day, UK
Fiorella De Cindio, Italy
Heidi Feng, USA
Anita Komlodi, USA
Piet A.M. Kommers, The Netherlands
Andrew Laghos, Cyprus
Stefanie Lindstaedt, Austria
Gabriele Meiselwitz, USA
Hideyuki Nakanishi, Japan
Anthony F. Norcio, USA
Ulrike Pfeil, UK
Elaine M. Raybourn, USA
Douglas Schuler, USA
Gilson Schwartz, Brazil
Laura Slaughter, Norway
Sergei Stafeev, Russia
Asimina Vasalou, UK
June Wei, USA
Haibin Zhu, Canada

Augmented Cognition

Program Chairs: Dylan D. Schmorrow, Cali M. Fidopiastis

Monique Beaudoin, USA
Chris Berka, USA
Joseph Cohn, USA
Martha E. Crosby, USA
Julie Drexler, USA
Ivy Estabrooke, USA
Chris Forsythe, USA
Wai Tat Fu, USA
Marc Grootjen, The Netherlands
Jefferson Grubb, USA
Santosh Mathan, USA
Rob Matthews, Australia
Dennis McBride, USA
Eric Muth, USA
Mark A. Neerincx, The Netherlands
Denise Nicholson, USA
Banu Onaral, USA
Kay Stanney, USA
Roy Stripling, USA
Rob Taylor, UK
Karl van Orden, USA

Digital Human Modeling

Program Chair: Vincent G. Duffy

Karim Abdel-Malek, USA
Giuseppe Andreoni, Italy
Thomas J. Armstrong, USA
Norman I. Badler, USA
Fethi Calisir, Turkey
Daniel Carruth, USA
Keith Case, UK
Julie Charland, Canada
Yaobin Chen, USA
Kathryn Cormican, Ireland
Daniel A. DeLaurentis, USA
Yingzi Du, USA
Okan Ersoy, USA
Enda Fallon, Ireland
Yan Fu, P.R. China
Afzal Godil, USA

Ravindra Goonetilleke, Hong Kong
Anand Gramopadhye, USA
Lars Hanson, Sweden
Pheng Ann Heng, Hong Kong
Bo Hoege, Germany
Hongwei Hsiao, USA
Tianzi Jiang, P.R. China
Nan Kong, USA
Steven A. Landry, USA
Kang Li, USA
Zhizhong Li, P.R. China
Tim Marler, USA
Ahmet F. Ozok, Turkey
Srinivas Peeta, USA
Sudhakar Rajulu, USA
Matthias Rötting, Germany
Matthew Reed, USA
Johan Stahre, Sweden
Mao-Jiun Wang, Taiwan
Xuguang Wang, France
Jingzhou (James) Yang, USA
Gulcin Yucel, Turkey
Tingshao Zhu, P.R. China

Human-Centered Design

Program Chair: Masaaki Kurosu

Julio Abascal, Spain
Simone Barbosa, Brazil
Tomas Berns, Sweden
Nigel Bevan, UK
Torkil Clemmensen, Denmark
Susan M. Dray, USA
Vanessa Evers, The Netherlands
Xiaolan Fu, P.R. China
Yasuhiro Horibe, Japan
Jason Huang, P.R. China
Minna Isomursu, Finland
Timo Jokela, Finland
Mitsuhiko Karashima, Japan
Tadashi Kobayashi, Japan
Seongil Lee, Korea
Kee Yong Lim, Singapore
Zhengjie Liu, P.R. China
Loïc Martínez-Normand, Spain
Monique Noirhomme-Fraiture, Belgium
Philippe Palanque, France
Annelise Mark Pejtersen, Denmark
Kerstin Röse, Germany
Dominique L. Scapin, France
Haruhiko Urokohara, Japan
Gerrit C. van der Veer, The Netherlands
Janet Wesson, South Africa
Toshiki Yamaoka, Japan
Kazuhiko Yamazaki, Japan
Silvia Zimmermann, Switzerland

Design, User Experience, and Usability

Program Chair: Aaron Marcus

Ronald Baecker, Canada
Barbara Ballard, USA
Konrad Baumann, Austria
Arne Berger, Germany
Randolph Bias, USA
Jamie Blustein, Canada
Ana Boa-Ventura, USA
Lorenzo Cantoni, Switzerland
Sameer Chavan, Korea
Wei Ding, USA
Maximilian Eibl, Germany
Zelda Harrison, USA

Rüdiger Heimgärtner, Germany
Brigitte Herrmann, Germany
Sabine Kabel-Eckes, USA
Kaleem Khan, Canada
Jonathan Kies, USA
Jon Kolko, USA
Helga Letowt-Vorbek, South Africa
James Lin, USA
Frazer McKimm, Ireland
Michael Renner, Switzerland
Christine Ronnewinkel, Germany
Elizabeth Rosenzweig, USA
Paul Sherman, USA
Ben Shneiderman, USA
Christian Sturm, Germany
Brian Sullivan, USA
Jaakko Villa, Finland
Michele Visciola, Italy
Susan Weinschenk, USA

HCI International 2013

The 15th International Conference on Human–Computer Interaction, HCI International 2013, will be held jointly with the affiliated conferences in the summer of 2013. It will cover a broad spectrum of themes related to human–computer interaction (HCI), including theoretical issues, methods, tools, processes and case studies in HCI design, as well as novel interaction techniques, interfaces and applications. The proceedings will be published by Springer. More information about the topics, as well as the venue and dates of the conference, will be announced through the HCI International Conference series website: http://www.hci-international.org/

General Chair
Professor Constantine Stephanidis
University of Crete and ICS-FORTH
Heraklion, Crete, Greece
Email: cs@ics.forth.gr

Table of Contents – Part I

Part I: Design and Development Methods and Tools

Part II: Information and User Interfaces Design

Part III: Visualisation Techniques and Applications

Part IV: Security and Privacy

Part V: Touch and Gesture Interfaces

Part VI: Adaptation and Personalisation

Part VII: Measuring and Recognising Human Behaviour

Table of Contents – Part II

Part I: Access to Information

Part II: Supporting Communication

Part III: Supporting Work, Collaboration, Decision-Making and Business

Part IV: Mobile and Ubiquitous Information

Part V: Information in Aviation

Part I

Design and Development Methods and Tools

Visual Programming of Location-Based Services

Antonio Bottaro[2], Enrico Marino[1], Franco Milicchio[1], Alberto Paoluzzi[1], Maurizio Rosina[2], and Federico Spini[1]

[1] CAD Laboratory, Dept. of Computer Science and Automation, Roma Tre University, Via della Vasca navale 79, 00146 Roma, Italy
[2] SOGEI S.p.A., Via M. Carucci 99, 00143 Roma, Italy
abottaro@sogei.it, enrico125@gmail.com, fmilicchio@me.com, apaoluzzi@me.com, mrosina@sogei.it, federico.spini@gmail.com

Abstract. In this paper we discuss a visual programming environment for design and rapid prototyping of web-based applications, securely connected to remote Location-Based Services. The visual programming approach of this research is based on computation as data transformation within a dataflow, and on visual composition of web services. The VisPro environment uses a very simple approach to service composition: (a) the developer takes a set of web widgets from a library, (b) builds interactively a user interface by drag and drop, (c) builds the application logic of the web service by drawing the connections between boxes (standing for suitable data transformations) and widgets (standing for user interaction). The development session produces, in presentation mode, a web page where the user may trigger, and interact with, the novel data mining and related computation. A successful GUI (and logic) is abstracted as a new service, characterized by a new widget, and stored in the widget library.

1 Introduction

Cartographic Applications, traditionally concerning projects related to digital cartography, like Geographic Information Systems, Digital Elevation Models, Terrain Analysis and Remote Sensing, have always required a great amount of vertical specialization, with reference to the used technologies, and to the professional skills needed to implement the applications. The new world of web radically changed this scenario; cartographic applications quickly moved from GIS (Geographic Information Systems) technology, currently downgraded to fill useful but peculiar application niches, towards Geographic Information Services. In few easy words: from GIS to ubiquitous GIS. Such new world of ubiquitous GIS is totally oriented to, and supported by, information services connected to geolocalization — i.e. Location Based Service. LBS provides advanced cartographical interfaces, directly at final user's disposal, and may convey easy customization of information services that, hinged around localization attributes, become flexible, interactive, and strongly integradable.

At the present time, web technology is going to make available software frameworks usable as SaaS (Software as a Service) components, that puts the user in control of programming— through suitable libraries (API/SDK)—highly specialized services, which are instantly deployable within the departments of a public institution or private

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 3–12, 2011.

company. The predictable future of this scenario, looking beyond the infrastructural components and the evolution of technological platforms, requires the adoption/development of novel visual programming environments, that must enforce the easiest (i.e. strongly based on drag-and-drop) composition of elementary programmable components, in order to build highly specialized applications and location-based services. This predictable trend is especially desirable when dealing with applications of location intelligence, where sophisticated techniques of business intelligence require a real-time visualization of discrete events or diffuse situations that may evolve dynamically at various geographical levels, i.e. either countrywide or within specified regional districts. Such analyses may concern also investiga- tion activities, in fraud prevention and hampering, that may require autonomous elaboration capability with high levels of secrecy.

Fig. 1. Example of LBS (Location-Based Service) developed by SOGEI, a company fully owned by Italian Ministry of Economy and Finance, for a related Agency

2 Background

The visual programming approach of this research is based on a few simple-but-pervasive concepts, namely (a) on computation as data transformation within a dataflow, and (b) on visual composition of web services that are going to shape the rapid application development in the next years, where on-demand enterprise IT services cry for a dynamically configurable architecture of the service process engine [1]. Whereas other approaches [5, 1] use either web reasoning or planning [1] or Petri nets [7] for service composition, as in Ref. [3] we use a simple but effective dataflow model. We briefly describe in the following the two fundamental abstractions that are used extensively in the project. Let us just recall that abstraction is the process of ignoring the details, and to focus on the overall design (the big picture).

2.1 Composition of Web Services

As defined by the W3C, a web service is a software system designed to support interoperability between different computers on the same network. The main characteristic of a web service is to provide a software interface that other systems can interact with. The interface is operated through messages included in an envelope. These messages are usually transported via the HTTP protocol and formatted according to the standard XML, using some lightweight protocol for exchange of messages between software components. Software applications, possibly written in different programming languages, and deployed on different hardware platforms, operate data exchange and execution of complex operations either on corporate networks or the Internet, via the interfaces they expose publicly and through the use of operations they make available. Web service-based architectures generally, but not necessarily, make use of XML data representations. More efficient implementations, like the one we discuss in this paper, use JSON. In particular, we discuss a novel approach that establishes an initial library of interoperable software components, including location-based services and related interfaces, as template GUI components, and accumulates successful business processes as templates for new service classes.

2.2 Computations as Data Transformations

The abstract concept of function provides an important computational abstraction, since it encapsulates the type of computation and hides the details of the calculation to the user. This one only needs to know the mechanism of the function call, and not how the function works. A single function superbly represents the essential characteristics of a computation. Graphically, a function may be represented by a rectangle (box) and by its function name. In abstract terms, it is a transformation (mapping) between the set of possible data (input) and the set of possible outcomes (output). In our visual programming approach, we only use two symbols, possibly instantiated in a set of specialized icons, to represent (at variable levels of detail) programs and data, respectively. Our computing environment supports higher-level functions, i.e. services (programs) that accept other services in input and/or that produce other services. All functions in this environment are unary, i.e. accept only one input and provide only one output. Were this condition is not satisfied, either data containers or curried higher-order functions are used. The first case is resolved visually with a single rounded rectangle that hides the detail, the second one with a composition of partial functions that are associated in a bijective manner to data components [9]. The whole service can be viewed as a stateless data-flow graph with pipelined data exchanges. In particular, we produce computable diagrams, which may be directly executed for debugging, or exported as concurrent processes, or encapsulated and abstracted as a new higher-level service component. Each message in each executable diagram can be stepwise graphically visualized on request.

2.3 A Righteous Mixup of Web Technologies

In this section we shortly discuss the selected set of web technologies that our visual programming approach is based upon. They include prototype-based functional programming with JavaScript, non-blocking Ajax methods, and Json-based data representations.

Client-side scripting via JavaScript. The JavaScript language is used to gain access to programmable objects in a client computational environment, typically a web browser. It is characterized as an object-oriented scripting language, dynamically typed and prototype-based. JavaScript is also considered a first-class functional language, since it contains both closures and higher-order functions. From this viewpoint, is it quite similar to Scheme, even if using a C-like syntax. JavaScript is mainly used as a client-side language implemented within a web browser, in order to offer powered user interfaces and access to dynamic web sites. The main principles of the language design are derived from Scheme, like lexical closures, i.e. namespaces of local symbols, and lambda functions, i.e. anonymous functions, to be used as input/output values to/from other functions, as well as from the Self language (prototype-based programming).

Prototype-based programming. Prototype-based programming is an object-oriented style without classes, where the reuse of behaviors—known as inheritance in object-orientation based on classes—is obtained through a cloning process of existing objects, that behave like prototypes, i.e. as archetypal examples. This programming model is also known as classless or prototype-oriented, and also as instance-based programming. The model contains two methods to build new objects: a literal form for “ex nihilo” creation, and another one via the cloning of existing objects, allowing for the addition of new properties non present in the prototype. This programming paradigm looks as the best one in order to accommodate the development of user interfaces that are both flexible and sophisticated.

AJAX methods. AJAX (standing for Asynchronous JavaScript and XML) is a group of technologies for web development, used client-side to create interactive web applications through non-blocking server calls using JavaScript. By using AJAX, web applications asynchronously receive data that are sent in background without interfering with the display and the behavior of the active pages. Furthermore, AJAX uses a combination of HTML and CSS to define the presentation style of the information. The DOM model of the web page is accessed via JavaScript to dynamically display the page, and to allow the user to interact with the information thereof. For this purpose, JavaScript and XMLHttpRequest supply asynchronous data exchange methods with the server, to avoid the browser to refresh the whole page. Recently, JSON (JavaScript Object Notation) and JavaScript are more and more utilized as an alternate efficient format for data exchange and as data manipulation language, respectively. In particular, the jQuery library supports a complete suite of AJAX methods. Objects and method contained thereof may open the browser multiple asynchronous data channels to the server without requiring a refreshing of the current page.

JSON data representation and exchange. JSON is a lightweight text format for data exchange. Since it is text-based, the data format is easy to read and to write for humans, and possesses a regular syntax that is easy to parse automatically. Even more, JSON is exactly a subset of JavaScript, whose parsing is even easier than the parsing of XML. JSON code is a literal representation of Javascript arrays and objects, that can be nested at will. In particular, JSON is built over efficient access to two universal data structures that are virtually supported, in one form or in another, by all modern programming languages:

1. Collection of name/value pairs: in other languages, it is implemented as an object, record, struct, dictionary, hash table, keyed list, or associative array.
2. Ordered list of values: in most languages it is implemented as array, vector, list, or sequence.

3. Such two fundamental structures can be mutually nested in any way, therefore allowing to easily represent any type of data structure.

HTML5 Canvas. The HTML5 standard provides advanced functionalities needed by novel Rich Internet Applications (RIA). In particular, the canvas object yields raster visualization of any complexity and precision, that previously was provided only by vector web graphics based on SVG (Scalable Vector Graphics), that needed a suitable plugin in the browser. The HTML5 document defines the <canvas> element as “a resolution-dependent bitmap canvas which can be used for rendering graphs, game graphics, or other visual images on the fly.” In other words, a canvas is a page rectangle were the application can draw by using JavaScript. HTML5 provides a set of functions (canvas API) to draw geometric shapes, define open or closed vectorial paths, create color gradients, and apply geometric transformations and image filters.

WebGL: advanced web graphics. The embedded JavaScript code writes the canvas DOM elements with drawing functions that may produce both 2D and 3D interactive graphics, even generated via the hardware support of the GPU, if present, using the WebGL framework. WebGL is a multi-platform and multivendor standard API for low-level 3D web graphics based on OpenGL ES 2.0 (the version for mobile devices), exposed through HTML5 Canvas as elements of the DOM interface. Several libraries are currently being developed to allow an higher-level employment of the amazing graphics power exposed via WebGL canvas and the JavaScript language directly in the web page by the browser, and without the use of any specialized plugin.

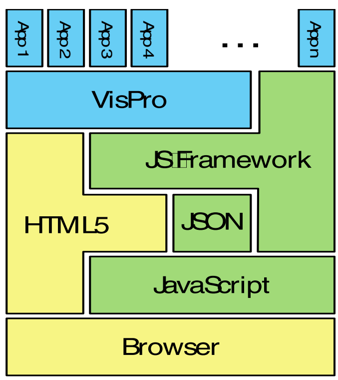


Fig. 2. A stack metamodel of the VisPro visual programming environment

3 User Interaction Methods

We distinguish here between two main classes of user-interaction methods, namely the development style, committed to the interactive generation of executable diagrams and user interfaces, and the navigation style, used for navigation of information diagrams and for exploration of the hierarchical structure of basic and complex widgets, as well as for the interactive exploration of the logic of rich internet applications. The development style, or 2D visual programming style, is based on drag-and-drop interaction, icon generation, input-output of widgets from a structured library of programmable objects. The navigation style, or 3D visual documentation style, presents the user with a dynamic view of the hierarchical structure of a rich internet application. The 3D style can be used for purposes of interactive software documentation and for training of developers of RIAs and UIs.

3.1 Visual Programming (2D)

The widgets (or controls) of the GUI (Graphical User Interface) are defined as reusable elements of the graphical user interface. They offer a pleasant layout of the information and provide the user standardized techniques for data manipulation. In particular, a widget is a GUI element that produces some data input/output in a way more or less modifiable by the user; for example just consider a GUI window or a text box.

The main characteristic of a widget is to give the user a single interaction point for the direct manipulation of a specific type of datum. In other words, the widgets of the user interface are single visual building blocks that, when suitably combined into an application, may visualize both the data elaborated by the application and, in variable measure, the relations between them and the application logic. Of course, a web widget is a software control for the web. In practice, it is a small application that can be installed within a web page and can be executed directly by the final user. A web control is often a stand-alone application that can be embedded from third-party sites by a user with suitable write access to somewhere (e.g. to a web page, a blog, the user profile in some multimedial social site, etc.).

The VisPro environment uses a very simple approach to service composition: (a) the developer takes a set of web widgets from a library, (b) builds interactively a user interface by drag-and-drop operations, (c) builds the application logic of the web service by drawing the connections between boxes (standing for suitable data transformations) and widgets (standing for user interaction). The development session produces, in presentation mode, a web page where the user may trigger, and interact with, the novel data mining and related computation. A successful Interface (and logic) is abstracted as a new web service, characterized by a new widget, and stored in the widget library.

4 The VisPro Architecture

Some preliminary definitions may be useful to compactly set out the main characteristics of the visual programming environment VisPro discussed here.

1. *logical workspace* (LW): work area where to define the application logic;
2. *visual normalized space* (VN): work area where to define the UI;
3. *databox*: basic element of logical workspace that represents a data object;
4. *funcbox*: basic element of logical workspace that represents a function (service or data transformation)
5. *inlet*: connection area for input links to a box (databox or funcbox);
6. *outlet*: connection area for output links from a box (databox or funcbox);
7. *link*: connection between the outlet of a box and the inlet of another box.

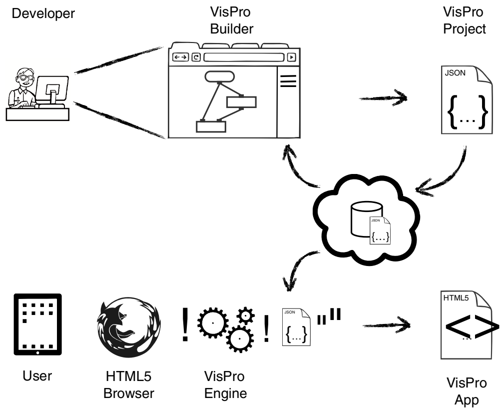


Fig. 3. The process of producing and using visual programming applications. The repository of widgets (one-to-one with data services) plays the central role.

4.1 Workspace Layout

According with the well-known design rule of information systems, that requires a strong separation between the application logic and the application presentation, the whole VisPro workspace is split into two adjacent areas denoted as LW and VN, respectively, that become alternating when the tool will be ported to mobile devices. The LW is responsible for the application logic definition, whereas the VN is in charge for the visual definition of the user interface of the new service, including the layout of the presentation and the positioning and dimensioning of the component widgets. A service layout typically possesses a hierarchical structure, being composed by widgets properly instanced and arranged by means of affine transformations (translations, rotations, scalings) within the Visual Normalized space. Since each widget may in turn be composed by others, the hierarchical structure of the VN content is represented as a directed acyclic multigraph (we refer to it as the presentation graph of the service). The VisPro interaction and building tools (single and multiple selection, transformation handles, etc.) allow the user to work in both the layout zones (LW and VN) at any chosen hierarchical levels, by selecting groups of widgets, as well as by using a stratified organization of elements by layers. Each object in the VN correspond to an object in the LW. The converse is not true. Some elements of the LW may have no mapping to elements in the user interface.

4.2 Computational Model

As already discussed in the Introduction section, the computational model underlying VisPro is the data flow model. Some properties neeed to be introduced to correctly represent a data flow graph.

Assumptions. In particular, we postulate the following assumptions:

1. a databox represents a datum and is provided with only one outlet;
2. a funcbox represents a function and is provided with one or more inlets and only one outlet;

3. to each inlet can be connected only one link;
4. multiples link can be conversely connected to the same outlet.

The visual programming approach is based on two main interaction operations:

1. the insertion of a box (funcbox or databox);
2. the linking of the outlet of a box to the inlet of another one.

Computational network. A funcbox produces an output dataflow by applying the function it represents to its input dataflows. The computation of a funcbox is only triggered when at least one datum is available for each funcbox inlet. The triggering of the function computation over a tuple of input data is sparked when the last needed datum is received by the funcbox. The triggering condition, together with the constraint concerning the existence of only one link for each inlet, allow to prove the nonexistence of flow loops in the computational network. The set of links determines a direct acyclic graph of the computation. The whole computation is modeled by a modified DFS (Depth First Search) traversal that requires that all the inputs necessary for the trigging of each single computation are available. The traversal ordering is established by visiting, in connection order, the first node with all input available.

*Example***.** Let the output of node A in Figure 4 be connected with an input of B (requiring two inputs) and with C (single input). Let the C output linked with the other input of B. At first, when the computation is started from A, the triggering of the B computation is attempted. But the tentative cannot proceed because the second datum of B is missing. Therefore the A node starts the C computation. When halted, C will attempt the triggering of B. This time the try has success because the B inputs are all available.

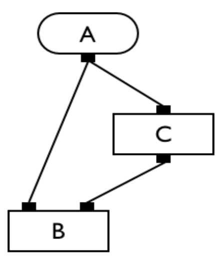


Fig. 4. Example of Warnier-Orr diagram

4.3 Process Communication

The linking of two boxes is always directed from an outlet towards an inlet, and no vice versa. In particular an inlet is connected to an outlet if interested to the flow of data passing through it. The action of connecting can be seen as a registration at the notification of an event:

1. the outlet is the *event publisher*;
2. the inlet is the *event subscriber*;
3. the release of a datum by the outlet is the *triggering event*;

When a databox releases an instance of its datum, or when a funcbox relase an instance of its output, the outlet generates the event associated to the release of the datum, and notifies all the inlets interested to such datum.

Process registration and notification. Let us recall that within the JavaScript language the functions are first-class objects, and that the scope is lexical, so that the functions have access to the context they are created within. This amazing language feature allow one to exploit the callback functions as connection apparatus. In particular, the registration consists in passing a callback function from the inlet to the outlet. Conversely, the notification consists in the execution by the outlet of all the callback functions registered thereof. Also, the callback function employed in the registration maintains a full access to the context where it was created (i.e. to its closure) so transmitting its own knowledge to its publisher, that does not use a copy of it, but the function itself.

Triggering the funcbox computation. Each inlet has an associated callback function. The execution of the callback by the outlet provides the flow of the datum from the outlet to the corresponding inlet. Therefore, when a funcbox receives an input datum (via the inlet), verifies that all the other inputs are present and, if this predicate is true, applies to them the function associated to it (that we can call also box behavior). When the execution terminates, or when an item of the output stream has been generated, the funcbox will execute the callback functions registered on its outlet.

Data caching. Let us consider a funcbox b with two inlets i1 and i2. Let a datum be arrived in i1 but not in i2. The b computation cannot start, of course. The arrival of a new datum in i1 would imply the loss of the previous one. Therefore, each inlet is provided of a queue temporary storing the transient dataflow. When a computation is triggered the needed data are retrieved from the queues. No data loss results from this very simple caching approach.

4.4 Service Security

In order to provide users with a secure and reliable web-based service, our system implements a standard two-phases authentication facility [12]. In our environment, the client ensures the authenticity of the server, as well as enforcing a secure communication, connecting through an SSL/TLS [13, 2] channel. User authentication is performed comparing hashed versions of the secret passphrase entered by the user, with a calculated one on the server side; hashing functions on both sides are chosen within the Secure Hashing Algorithm family [6, 5], providing a high security level with respect to current cryptanalysis evaluations [8, 4]. An additional time-based control is performed, inspired by the Kerberos protocol [10], in order to ensure the authenticity of an authentication sequence within a man in the middle attack scenario: in this case, a challenge-response sequence requiring a long timeframe is a strong indication that the original message has been diverted, and the communication between the server and the (supposed) client cannot be considered secure. Once the identity of a user is established, it is associated with one or more group memberships, mimicking standard POSIX user-group links. Additionally, each object possesses an associated access control list (ACL), reflecting standard POSIX.1e capabilities [11], allowing our users to read, and modify, each object in our system with an object-access granularity.

5 Conclusion

The VisPro programming environment is currently under development using WebGL, HTML5 and a set of Javascript frameworks. We are going in few months to be able to make the first experiments of localization intelligence according to the strong initial requirements of this project.

References

1. Agarwal, V., Dasgupta, K., Karnik, N., Kumar, A., Kundu, A., Mittal, S., Srivastava, B.: A service creation environment based on end to end composition of web services. In: WWW 2005: Proceedings of the 14th International Conference on World Wide Web, pp. 128–137. ACM, New York (2005)
2. Allen, C., et al.: The transport layer security (TLS) protocol. RFC 5246 (August 2008)
3. Curbera, F., Duftler, M., Khalaf, R., Lovell, D.: Bite: Workflow composition for the web. In: Krämer, B.J., Lin, K.-J., Narasimhan, P. (eds.) ICSOC 2007. LNCS, vol. 4749, pp. 94–106. Springer, Heidelberg (2007)
4. De Cannière, C., Rechberger, C.: Finding SHA-1 characteristics: General results and applications. In: Lai, X., Chen, K. (eds.) ASIACRYPT 2006. LNCS, vol. 4284, pp. 1–20. Springer, Heidelberg (2006)
5. Eastlake, D.E., Hansen, T.: US secure hash algorithms (SHA and HMAC-SHA). RFC 4634 (July 2006)
6. Eastlake, D.E., Jones, P.E.: US secure hash algorithm 1 (SHA1). RFC 3174, Visual Programming of Location-Based Services 11 (September 2001)
7. Gao, M., Wu, Z.: EPN-based web service composition approach. In: Liu, W., Luo, X., Wang, F.L., Lei, J. (eds.) WISM 2009. LNCS, vol. 5854, pp. 345–354. Springer, Heidelberg (2009)
8. Matusiewicz, K., Pieprzyk, J., Pramstaller, N., Rechberger, C., Rijmen, V.: Analysis of simplified variants of SHA-256. In: Western European Workshop on Research in Cryptology (2005)
9. Milicchio, F., Bertoli, C., Paoluzzi, A.: A visual approach to geometric programming. Computer-Aided Design and Applications 2(1-4), 411–421 (2005)
10. Neuman, B.C., Tsó, T.: Kerberos: An authentication service for computer networks. IEEE Communications 32(9) (1994)
11. IEEE. IEEE standards interpretations for IEEE standard portable operating system interface for computer environments. IEEE 1003.1-1988/INT (1992)
12. Stallings, W.: Network Security Essentials, Applications and Standards. Prentice-Hall, Englewood Cliffs (2000)
13. Wagner, D., Schneier, B.: Analysis of the SSL 3.0 protocol. In: USENIX Workshop on Electronic Commerce (1996)

Connecting Envisioning Process to User Interface Design Process

Naotake Hirasawa[1], Shinya Ogata[2], and Kiko Yamada-Kawai[1]

[1] Otaru University of Commerce, Midori 3-5-21, Otaru, Japan 047-8501
{hirasawa,kawai}@res.otaru-uc.ac.jp
[2] U'eyes design Inc., Housquare Yokohama 4F, Nakagawa 1-4-1, Tsuzuki-ku, Japan 224-001
ogata@ueyesdesign.co.jp

Abstract. As embedded systems have increasingly become multifunctional, it has become necessary to clarify what sorts of experiences the user can have through these functions. It is the system's user interface (UI) that leads this user experience, therefore, UI design has increased in importance. However, conventional design processes, often set the UI design after defining the system requirements, which causes operations to become more complex and leads to a lack of operational consistency. To overcome these problems, a design process was conceived whereby the user experience envisioned in the product planning phase is realized in the UI. Two workshops were held for development-related personnel to introduce the conceived process into actual development projects. The effectiveness of the process was acknowledged by all the participants, but they also pointed out the difficulty of incorporating the process into their organizations.

Keywords: User Interface Design, Design Processes, Workshop.

1 Introduction

In developing embedded systems, the importance of the UI design has been increasingly recognized. The means of operating the system functions are realized into the UI in the embedded system. When the functions are few, the UI is simple and easy to operate. A rapid advance in infrastructure technology, however, has enabled the system to incorporate a multitude of functions; consequently, UIs have become complicated. Systems now face the contradiction that convenient functions are incorporated but they cannot be fully utilized as the operation becomes too complicated.

At the same time, the linkage with telecommunications technology has enabled a system to connect to the information infrastructure and other systems, providing it with intermediary roles for various services along with services of its own. The cell phone is a typical example as it started from a single phone unit to evolve into an information terminal that now intermediates various other services. This means a UI needs to be considered at the phase of development where the services to be provided by the products are conceived. In other words, the UI must start to be considered in the product planning phase.

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 13–19, 2011.

Taking into consideration these changes in the role of the UI in an embedded system, our research project has created a design process that allows the user experiences (UX) envisioned in the planning phase to be realized into the UI, while balancing the operational system as a whole.

This paper describes a summary of this process model, and discusses the issues involved in the organizational implementation of the process, which were based on the analysis of the results of workshops held for company employees engaged in development.

2 Process Model

2.1 Problems of UI Design in Embedded System Development

In general, UI design proceeds based on the functional and non-functional requirements that are clarified after the system requirement definition have been conducted. In this process, time is so limited by software design that the UI design is often implemented in the system without performing any validation from the user's perspective. Even when UI design is outsourced, UI specification is often designed under conditions where all the functions to be implemented have already been determined. In such cases, the following problems will arise in designing the UI during the embedded system development [1].

- There is a possibility that the UXs envisioned in the planning process are not necessarily reflected in the UI.
- The direction of the multifunctionality of the system is not controlled, so operational systems become complicated.
- Validation of the UI specification becomes difficult.
- Management in changing the UI specification becomes difficult.
- To design the optimal UI specification, user researches conducted in the planning phase sometimes need to be performed again.

To overcome these problems, such measures are available as (i) coordinating the requirements for defining the embedded system and for designing the UI; and (ii) making the background of the UI specification traceable. Besides these, our project implemented the following two new measures.

1) Introduction of the concept "view for context of use"

The performance of designing UI over the past 20 years at U'eyes Design Inc. clearly shows that the defining *view* that characterize activities of the users of the system can facilitate UI design [2]. For example, if we consider the UI for displaying the sales performance of a POS system, the UI can be designed from at least three different viewpoints: from the system's goal of checking the sales performance; from the viewpoint of where the system is used; or from the viewpoint of the user characteristics. The important thing is that these viewpoints determine the directions of UX when the system is used. In other words, clarifying the viewpoints will also lead to the identification and clarification of how the system will appeal to the user.

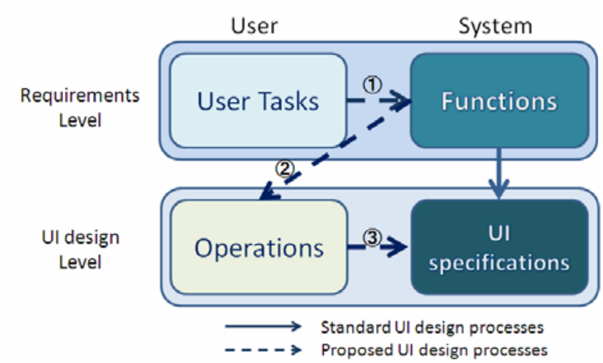


Fig. 1. Depicting a framework of UI design processes

If the concept of *view* is used, user characteristics will be the viewpoint in the method advocated by Cooper [3]; and it can be regarded that UX is defined from the viewpoint of “work” in Contextual Design [4].

Defining these viewpoints means clarifying the intentions or directions of using the system, and this becomes an aid to determining the direction of the UI during the planning phase. We have named these views as “*view for context of use (riyou tokusei in Japanese)*” as they determine the nature of the system use.

2) Depicting the concept of proposed UI design processes

The general rule for designing the UI in the embedded system is to design the UI specification based on the functions determined by the definition of the requirements. In this case, the UI design process is people-dependent and the UI is often designed with reference to the conventional UI or a competitor’s UI.

We determined the reason for such a practice to be that the basic concepts underlying the UI design process was not shared among members. This was because the background the UI was designed against was ambiguous. To help solve the problem, the following basic concepts were identified regarding the relationship between the UI and the system requirements: (Fig.1)

- The goal of the system functions is to realize user tasks.
- The system functions are used through the user’s operation of the system.
- The means to realize the operation of the system is the UI.

If the design process is organized under these concepts, UI design can be positioned as the means to realize user tasks. As a result, UI specification can be traced from the user tasks.

2.2 Proposed Design Processes Model

Based on the above-mentioned ideas, it becomes possible to develop a process model that can connect activities from planning to UI design. In developing the process, some development lifecycle models were used: Embedded System development Process Reference (ESPR) developed by Software Engineering Center under Information-Technology Promotion Agency was used as a base [5]; and ISO 15288

[6] and ISO 12207 [7] were used as reference. Also ISO 13407 [8], ISO 18529 [9], Rapid Contextual Design [4], IBM's UCD model [10], and Cooper's model [3] among others were used, especially for process models during the UI design process (Fig.2).

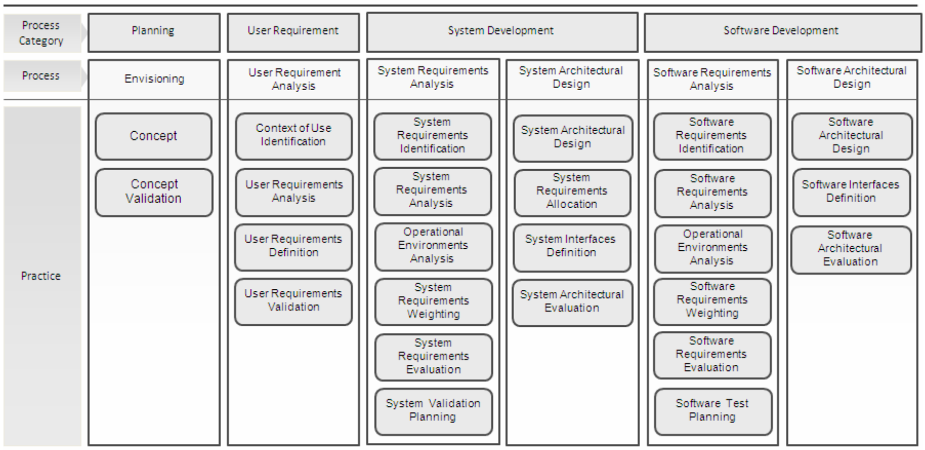


Fig. 2. Outline of HCD Process integration Model

3 Implementing Workshops Based on Proposed Design Process

To confirm the validity of the proposed process model, workshops were held for employees engaged in the development in domestic manufacturers, to investigate its possibilities and limitations. Workshops were implemented at the following two companies:

3.1 Case A (Manufacturing Equipment Maker)

Background. A large manufacturing equipment maker with over 10,000 employees, this company has created a brand with high quality products and customer-oriented services. The company is aiming to convert its business from providing manufacturing equipment to providing solutions for controlling multiple equipments. In order to fulfill the goal, a new design process needed to be developed. Workshops were held with the expectation that the process model we had conceived would be applied to their new design process.

Participants. 5 engineers of equipment development and software development, and 1 sales-related employee.

Program. A series of four workshops were held from August 2009 to December 2009. Each goal of the workshop was the following;

1st was to understand new requirements definition processes,
2nd was to identify as-is scenarios for describing contexts of system use,

3rd was to elicit system requirements,
4th was to consolidate and depict system requirements.

In the program, a solution that had been already developed was used as the object, and various activities were performed from the planning phase through to the system requirements analysis based on the proposed design process. Information necessary for the workshops was searched for and collected before the workshops began. Eventually a definition of the system requirements was established. After all the workshops finished, the participants were asked to evaluate the proposed processes with questionnaire and interviews were conducted for the every participant.

Results. The system requirements analysis was able to clarify the functions that the existing solution lacked. All the participants acknowledged the effectiveness of the proposed design process. Specifically, participants pointed out the possibility of sharing information from the planning phase to the lower level processes. One participant commented that the process provided a bird's-eye view over the upstream processes of the system development, and that therefore, it could be directly applied to educational training.

In terms of the application possibilities of this process to their actual operation, three participants (50%) said it could be applied. Similarly, three participants said they could immediately apply the process to their own work. Half of the participants pointed out the difficulty of expanding the process to the entire organization. Some expressed anxiety over the possibility of implementing the process by themselves after the workshops were over.

3.2 Case B (Home Appliance Maker)

Background. This company is a home appliance maker with a long history and more than 10,000 employees. It has state-of-the-art technologies in some specific technological fields, and has an established brand. However, the market to which the company provides products has matured, and differentiation of products between competitors has become difficult. The company has aimed at differentiation by making the UI of its products more appealing, and it has sought a new UI design process to accomplish this goal. They took this workshop as an opportunity to review their UI design process to develop new products, since UI technology infrastructure was experiencing drastic changes, and they asked us to hold workshops for employees who were engaged in development.

Participants. 2 product planners, 1 product designer, 2 requirements engineers, 2 software engineers, 4 usability engineers (total of 11).

Program. A series of four workshops were held from July 2009 to December 2009. Each goal of the workshop was the following;

1st was to identify the problems of the existing UI design processes,
2nd was to identify the user requirements of a new product,
3rd was to identify the system requirements,
4th was to design the outline of UI specifications.

Before the workshops, a seminar was prepared to help the participants understand the new design process. The program started by conducting an analysis of the existing design process, then it identified the issues involved in the UI design process, and finally it established the outline specification of the UI of the product under development. Workshops, held in parallel with the product development, were held during the program, adjusted in line with the product development schedule. After the every workshops finished, interviews for the every participant were conducted.

Results. During this program various ideas not obtained from the concurrent design process were obtained. All the participants acknowledged the effectiveness of the proposed design process. Specifically, the program enabled them to give a meaning to the newly derived requirement functions, and to strengthen the product proposal. The applicability of the process to their work was also acknowledged by all. However, more than 30% of the participants stated that only some individual methods could be applicable, and that the introduction of the process would be difficult. Continual support to usability engineers was mentioned as one important condition for applicability.

Both companies clearly acknowledged the effectiveness of the proposed process. As for the reason for the difference in applicability between the two companies, it can be pointed out that the former company does not offer organized support on human-centered design.

Whether trying to increase the conceptual power for solutions, or trying to enhance the appeal to the product by using UI, it has become clear that there is a big barrier to changing the existing design process, although the effectiveness of the proposed process was acknowledged. One of the issues is how to organize a support system for the proposed process.

4 Conclusions

As the embedded system has become increasingly multifunctional, the UI of the system has attracted expectations from the perspective of product appeal, even though it has a number of issues, and the importance of the UI design has increased. However, in the conventional design process, UI design is often conducted in the post-process of defining the system requirements, during the time the operational methods of the required functions are being realized. As a result, many design processes still have such problems as product specification turning out to be filled with functions, and the difficulty of changing the UI specification.

This study proposed a design process that enables the UX envisioned in the product planning phase to be implemented into the system in the form of a UI while balancing the operational system of the entire system. What characterizes this process is the concept of "views for context of use," and the direction given to the process based on the relationship between user tasks and functions.

Workshops were held for development personnel to check the validity of the proposed process model. As a result, the process was acknowledged to be effective, however, participants also pointed out the difficulty of actually introducing the process into their organizations.

Further study is expected on the process maturity model when the process has been introduced into organizations.

Acknowledgements. The authors would like to thank Toshiyuki Oketani and Hideaki Kasai for coordinating the workshops.

References

1. Hirasawa, N., Ogata, S., Urokobara, H.: Substantializing User Experience in the Processes for Developing Embedded Systems (in Japanese). Econimic Review 60(4) (2010)
2. Ogata, S., Urokobara, H., Aoshima, H.: Possibility of Modeling for Incorporating UI Design into System Design (in Japanese). In: Proceeding of Human Interface Symposium 2004, pp. 1073–1076 (2004)
3. Cooper, A., Reimann, R., Cronin, D.: About Face 3: The Essentials of Interaction Design. John Wiley & Sons Inc., Chichester (2007)
4. Beyer, H., Holtzblatt, K., Wood, S.: Rapid Contextual Design: A How-to Guide to Key Techniques for User-Centered Design. Morgan Kaufmann, San Francisco (2004)
5. Software Engineering Center 2007 Developing Process Guide for Embedded Software (in Japanese). Information-technology Promotion Agency, Japan, Syou-Ei-sya (2007)
6. ISO/IEC 15288:2008 Systems and software engineering - System life cycle processes
7. ISO/IEC 12207:2008 Systems and software engineering - Software life cycle processes
8. ISO 13407:1999 Human-centred design processes for interactive systems
9. ISO/TR 18529:2000 Ergonomics - Ergonomics of human-system interaction - Human-centred lifecycle process descriptions
10. Vredenburg, K., Isensee, S., Righi, C.: User-Centered Design: An Integrated Approach. Prentice-Hall, Englewood Cliffs (2001)

Learner-Centered Methodology for Designing and Developing Multimedia Simulation for Biology Education

Chi-Cheng Lin[1], Mark Bergland[2], and Karen Klyczek[2]

[1] Department of Computer Science, Winona State University, Winona, MN 55987, United States
clin@winona.edu
[2] Department of Biology, University of Wisconsin – River Falls, River Falls, WI 54022, United States
{mark.b.bergland,karen.k.klyczek}@uwrf.edu

Abstract. Biology students need exposure to modern research techniques relatively early in their educational careers. Computer multimedia simulation tools have been developed to address the challenge of providing all students with hands-on laboratory research experience. This paper presents a learner-centered approach to the design and development of a multimedia simulation for biology education. We present our methodology and a multimedia simulation tool designed and developed using the methodology. Out tool has been widely adopted by biological science educators for teaching molecular biology subjects in a wide range of undergraduate biology courses. We believe that our methodology can be adopted or adapted by learner communities in other disciplines.

Keywords: learner-centered methodology, case-based pedagogy, multimedia simulation, biology education, bioinformatics, workflow analysis.

1 Introduction

Early exposure to research is a national priority in science education: "All students should be encouraged to pursue independent research as early as is practical in their education" [1]. Therefore, biology students need exposure to modern research techniques relatively early in their educational careers. However, it can be logistically challenging to provide all students with hands-on laboratory research experience. Therefore, it is important to supplement wet laboratories with computer simulations to effectively engage students in the research process.

Numerous computer multimedia simulation tools have been developed with the attempt to address this challenge [2, 3]. Nonetheless, most of them have focused on the functional aspects and contents, with less attention being given to the mechanism to engage their users. For example, a tool may simply play animations or video clips for students to watch passively, providing little if any interactivity for the user [4, 5]. They contribute in explaining biological concepts such as the DNA transcription and sequencing technologies in molecular biology with a vivid visual aid, but they do not engage students actively. Recently, besides its role in multimedia simulation, computer

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 20–29, 2011.

software has become indispensable for scientific inquiries done by biologists on a daily basis [6]. The complexity and user-friendliness issues of the software tools have become an obstacle to effective research [6].

We believe that the design and development of an effective educational multimedia tool for studying biology should take a learner-centered approach. To do so, we need to answer the following three questions: 1) Does the tool effectively engage students using the tool? 2) Does the tool provide a user-friendly interface? 3) Does the tool effectively deliver contents essential to biology education?

The methodology we devised to answer these three questions is as follows: In the past we have used role-playing and simulations to enhance student interest in science, through cases emphasizing diagnosis and ethical issues surrounding health counseling [7]. Our latest innovation is to use this case-based pedagogy to actively engage students by placing them in virtual research settings, fostering their ability to solve problems confronted by research scientists [8, 9]. We analyzed the case workflow common in biological science investigations and designed a uniform graphical user interface suitable for biology education. The user interface also includes options for using the multimedia simulation in various ways to serve students at multiple learning levels as well as biology courses with different learning objectives.

We argue that the contents delivered by an effective biology education tool should stay current with the advances in the field of biological science. For instance, bioinformatics techniques have become fundamental to modern biological investigation [10, 11]. Students at multiple learning levels need to be aware of these techniques and competent in their use [12, 13]. However, because of their complexity, existing bioinformatics tools are learned most successfully when expert guidance is available [14, 15]. Moreover, laboratory techniques are commonly used to generate data for bioinformatics analyses. Many schools do not have the resources to provide hands-on experience in these techniques, especially for introductory biology students. Our approach is to integrate our laboratory simulations with bioinformatics tools commonly used to study fundamental concepts of bioinformatics. We analyzed the workflow of these bioinformatics tools to reduce their complexity, with the goal of making bioinformatics easier for its learners to understand.

The paper is organized as follows. Section 2 presents the overview of *Case It! v6*, the tool we designed and developed using our learner-centered methodology. Section 3 describes the user interface design. A scenario demonstrating how the student interacts with our tool to perform biology science investigation is given in Section 4. Section 5 concludes the paper and discusses future work.

2 Overview of Case It! v6

Case It! v6 is a tool that we designed and developed using our learner-centered methodology. The tool uses computer simulations to provide a context for student understanding by seamlessly integrating laboratory procedures with bioinformatics tools, using a case-based approach emphasizing important problems in biomedical research. It uses case-based pedagogy with an interactive virtual laboratory to engage students. Fig. 1 illustrates the opening screen of Case It v6. The devices and equipment shown

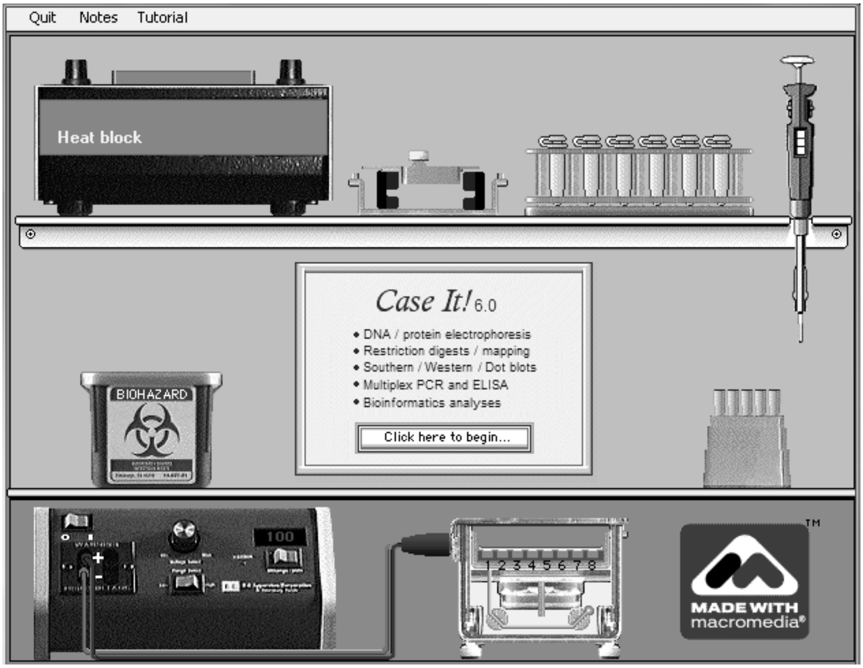


Fig. 1. Opening screen of Case It! v6

in the opening screen can all be operated to perform standard laboratory procedures virtually through multimedia interactions.

Case It! v6 consists of three major components. The schematic view of Case It! v6 architecture is depicted in Figure 2. The first component is the *Student Resource Manual*. The resource manual is a Web-based component that functions as an electronic resource manual. It helps students acquire background information about their cases by providing case descriptions and instructions for analyzing biological sequences associated with these cases. The second component is the *Case It Launch Pad*. This component is a collaboration environment that consists of a web page editor and bulletin board system. Virtual poster sessions can then be held and facilitated using bulletin boards associated with the posters.

The third component of Case It! v6 is the multimedia simulation, *Case It Simulator*. The simulator consists of two main functional units, namely the graphical user interface (GUI) and the functional core. The simulator provides a uniform graphical user interface to the students and allows them to play an active role in running biology wet labs on computers to investigate cases. The process can be visualized via interactive open-ended computer simulations that work with any DNA or protein sequence, and the results can be represented either in either textual or graphical format for analysis.

The functional core of the simulator consists of three modules, namely *Case Management*, *Simulation Control*, and *Extension/Integration*. The Case Management module organizes cases into self-contained folders with one case in each folder.

A case folder typically contains DNA or protein sequences for investigation. The Simulation Control module is the kernel of the entire tool, and is responsible for computation, programming logic, and multimedia simulation control. It interacts with the GUI unit and also bridges between the GUI unit and the Extension/Integration module. The Extension/Integration module is in place to address the needs of developing new biological tools to keep current with the advances in biological science. When a new concept such as bioinformatics has become essential for introductory biology students, we will either integrate it with existing cases or create new cases for it. Tools useful for delivering and reinforcing the new concept will then be explored and evaluated. As plenty of excellent tools developed by the field experts are freely accessible in the public domain, our strategy is to integrate suitable tools as “extensions” of the Case It! software instead of attempting to reinvent the wheels. Similar approaches have been adopted by the development of several integrated biological databases and analysis systems such as Biology Workbench [16] and SWAMI [17]. With this approach, the Extension/Integration module can focus on the integration mechanism to make the external tools intuitive and easy to use.

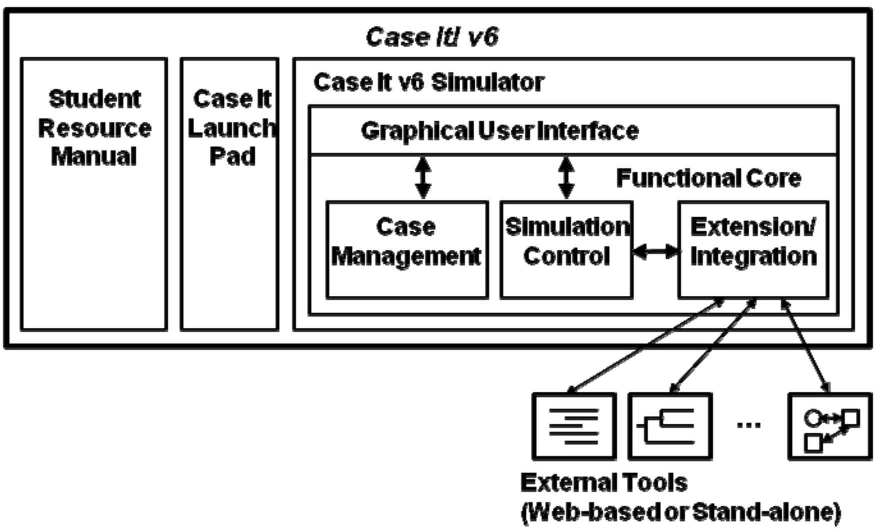


Fig. 2. Schematic view of Case It! v6 architecture

3 User Interface Design

We designed our user interface with the objectives to support our learner-centered, case-based methodology. The first objective aims at providing the learners with a consistent user experience in using the simulator. The second objective is aimed to ensure that the use of external tools is intuitive and effective. These objectives have been identified among key themes in the “new biologist-centric paradigm” for bioinformatics software design and development [6]. We analyzed the workflow of case-based biology science investigation and the workflow of the external tool to achieve these objectives.

2.1 Case Workflow Analysis

We believe that by providing a consistent user experience to the students, they can use the tool to study a wide variety of biology subjects more effectively. To achieve this objective, we began by analyzing the case workflow common in biological science investigations. The steps in the workflow is described as follows: 1) open a case in the context of the subject to be studied; 2) study the case description and its background information; 3) identify the problems to solve; 4) collect DNA or protein sequences as well as the enzyme that may be used to digest the sequences; 5) process the sequences using standard laboratory procedures, and 6) analyze the sequences using biotechnology or bioinformatics tools. The case workflow is summarized in Figure 3.

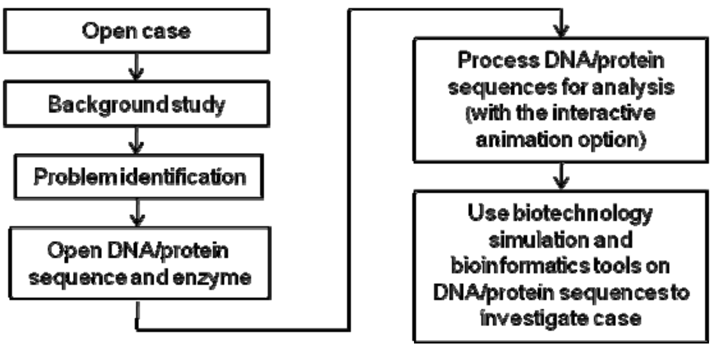


Fig. 3. Case workflow analysis for biology science investigation

Based on the case workflow, we designed a uniform graphical user interface for students to carry out the steps in the workflow. For example, students can open the DNA sequences, load them to the workspace of the simulator, and process them using virtual wet lab equipment. Moreover, we take students learning levels and biology courses learning objectives into consideration to make the simulator suitable for engaging a broader audience in biology education. As a result, the user interface of Case It! includes options for using the multimedia simulation in various ways. For example, students have the option to conduct wet labs virtually using interactive animations. On the other hand, students with more experiences in these procedures may choose to skip the multimedia interactions to proceed to the analysis of DNA data.

2.2 Extension Tool Workflow Analysis

As bioinformatics techniques have become fundamental to modern biological investigation, students at multiple learning levels need to be aware of these techniques and learn how to use them. The goal of the integration of our laboratory simulations with bioinformatics tools is to make bioinformatics easier for its learners to understand.

We began by identifying bioinformatics tools commonly used to study fundamental concepts of bioinformatics such as phylogenetic analysis and biological database search. We then analyzed the workflow of these bioinformatics tools to reduce their complexity.

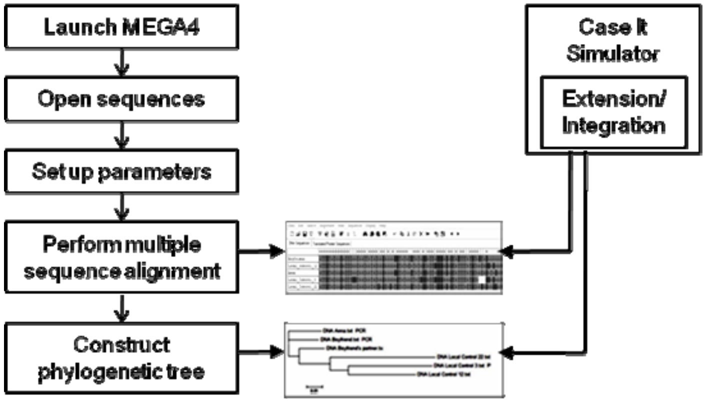


Fig. 4. Contrast between invoking MEGA4 (depicted in the left) and using Case It Simulator (depicted in the right) to perform multiple sequence alignment and construct phylogenetic tree

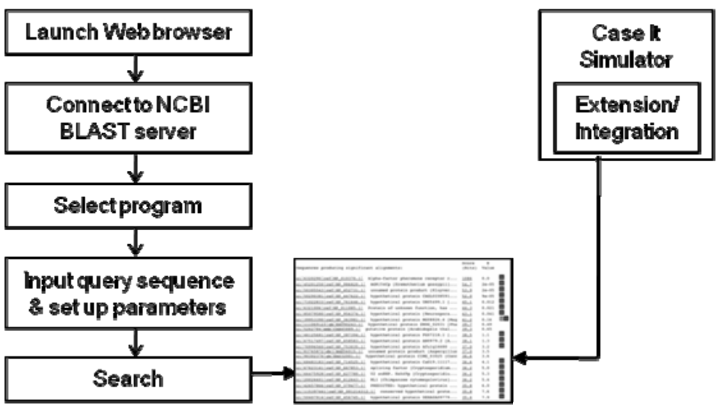


Fig. 5. Contrast between connecting to NCBI BLAST server (depicted in the left) and using Case It Simulator (depicted in the right) to carry out BLAST databases search

We adopted MEGA4 [18] for performing phylogenetic analysis and NCBI BLAST [19] for biological database search. The workflow of running MEGA4 directly requires the following steps: 1) open sequences; 2) set up parameters such as DNA or protein; 3) perform multiple sequence alignment; and 4) use the alignment to construct the phylogenetic tree. Our experiments show that it may take up to fourteen computer mouse clicks to display the alignment and the tree. Besides, students need to learn and understand how to set up parameters to use the tool correctly and effectively. Using the result of the workflow analysis, we designed and developed code to integrate MEGA4 with Case It! that only takes one mouse click to display the alignment and tree. The simplicity of our approach is illustrated in Figure 4, showing how the integration of Case It and MEGA4 greatly simplifies phylogenetic analysis.

Based on our study, the steps for searching the BLAST databases include: 1) connect to the BLAST Web server; 2) select the search program; 3) enter the query sequence; 4) set up parameters; and 5) search. Using the analysis, we designed and developed code to automate the workflow. The contrast between using a Web browser and using Case It! for BLAST database search is illustrated in Figure 5.

We believe that our approach allows students to focus more on biology science investigations than on learning bioinformatics tools, sometimes an overwhelming experience for beginning biology students. To enrich students at more advanced levels, Case It! also provides options for students to invoke bioinformatics tools directly. With these options, students can gain exposure to how those tools are employed by biological researchers in the real world.

4 Scenario

In this section we present a scenario to demonstrate how the student interacts with the Case It Simulator to carry out case-based biology science investigations. The case used in this example is regarding the subject of human immunodeficiency virus (HIV).

The student begins by studying the background information of the case and HIV, followed by identifying the problems to be investigated. After launching the Case It Simulator, the student opens the case folder and selects the DNA sequences to be analyzed and the enzyme to digest the sequences. The next step is to prepare the sequences for further analysis. At this stage, the students can choose to either load the digested sequences to the simulator's workspace automatically or carry out the wet

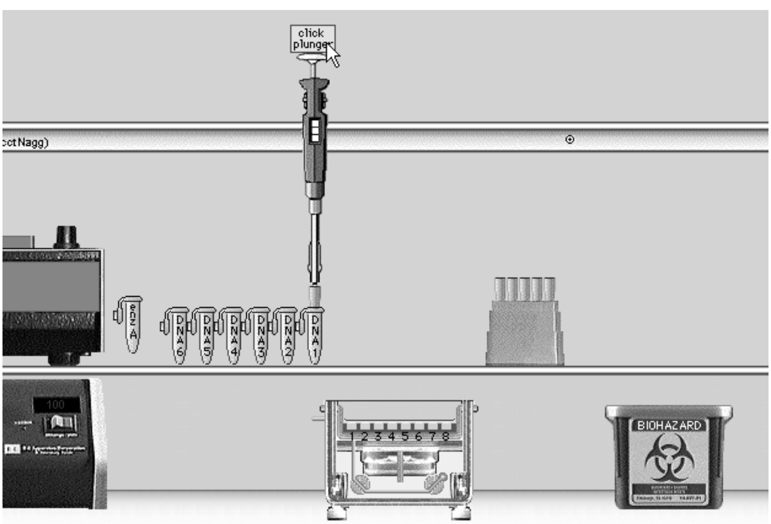


Fig. 6. Interactive multimedia simulation enables students to carry out wet labs virtually

Fig. 7. Screenshot of phylogenetic for the HIV case using Case It!

labs virtually using the interactive multimedia simulation. As shown in Figure 6, the student can use a pipette to obtain DNA sequences from the tubes and use them to perform standard wet lab procedures such as gel electrophoresis.

The student then starts the research and investigation after the sequence data are ready for analysis. If the student decides to investigate the evolutionary relationships among the HIV virus strands of people in this case, the student will perform phylogenetic analyses on the DNA sequences. The student can obtain the multiple sequence alignment and the phylogenetic tree with one simple mouse click from a pull-down menu in the Case It Simulator as shown in Figures 7. Instead of learning how to use MEGA4 first, the student can focus on the investigations using the alignment and the tree. The student then searches biological databases for similar sequences by using the integrated BLAST function in Case It!. This allows the student to concentrate on analyzing the results returned by the BLAST database server.

5 Conclusions and Future Work

In this paper we present Case It!, an open-ended, case-based multimedia simulation tool for biology education. We used a learner-centered methodology to design and develop the tool so that it would effectively engage students, provide a user-friendly interface, and effectively deliver current contents essential to biology education. Case It! has been widely adopted by biological science educators for teaching molecular biology subjects in a wide range of undergraduate biology courses [20]. Student surveys have been conducted to assess the effectiveness of the tool, and the result shows that the tool has improved student performance in learning molecular biology subjects [20, 21].

We are currently exploring the utilization of the Extensible Markup Language (XML) to convey biological data in various representations to serve users at multiple learning levels more effectively. We are also developing cases for other important molecular biology research tools such as proteomics and microarrays, and plan to integrate them with Case It! using our methodology.

Although our learner-centered methodology is applied to the design and development of a multimedia simulation for biology education, we believe that the same approach can also be adopted or adapted by learner communities in other disciplines.

Acknowledgments. The work presented in this paper is supported in part by NSF TUES (formerly CCLI) Program Grants DUE 0229156 and DUE 071577. We would also like to thank collaborators Mary Lundeberg, Arlin Toro, Rafael Tosado-Acevedo, C. Dinitra White, and Bjorn Wolter.

References

1. National Research Council, Committee on Undergraduate Biology Education to Prepare Research Scientists for the 21st Century: BIO2010: Transforming undergraduate education for future research biologists. National Academies Press (2003)
2. Bio-Soft Net, Education Section, `http://en.bio-soft.net/biocai.html`
3. Cybertory.org, `http://www.cybertory.org/`
4. Biology Animation Library, Cold Spring Harbor Laboratory's Dolan DNA Learning Center, `http://www.dnalc.org/resources/animations/`
5. Wellcome Trust Education Resources' Animations, `http://www.wellcome.ac.uk/Education-resources/Teaching-and-education/Animations/index.htm`
6. Kumar, S., Dudley, J.: Bioinformatics software for biologists in the genomics era. Bioinformatics 23(14), 1713–1717 (2007)
7. Bergland, M., Lundeberg, M.A., Klyczek, K., Sweet, J., Emmons, J., Martin, C., Marsh, K., Werner, J., Jarvis-Uetz, M.: Exploring Biotechnology Using Case-based Multimedia. American Biology Teacher 2006 68(2), 77–82 (2006)
8. Bergland, M., Klyczek, K., Lin, C.: Case It v6 (Programming by Bergland and Lin, case development by Klyczek) (2010), available for download, `http://caseit.uwrf.edu`
9. Klyczek, K.: Student Resource Manual for Case It! Version 6 (2010), `http://caseit.uwrf.edu/RM2010.html`
10. Butler, D.: Are you ready for the revolution? Nature 409, 758–760 (2001)
11. Gross, L.: Education for a biocomplex future. Science 288(5467), 807 (2000)
12. Gabric, K.: Bioinformatics in the biology classroom. Actionbioscience.org (2003)
13. Porter, S., Smith, T.: Bioinformatics in the community college. Journal of Industrial Microbiology 24, 314–318 (2000)
14. Flannery, M.: The Two Ts: Teaching and Technology. The American Biology Teacher 66(7), 506–510 (2004)
15. Lin, C., Zhang, M., Beck, B., Olsen, G.: Embedding computer science concepts in K-12 science curricula. In: Proceedings of the 40th ACM technical symposium on Computer science education (SIGCSE 2009), pp. 539–543. ACM Press, New York (2009)
16. Subramaniam, S.: The Biology Workbench – A Seamless Database and Analysis Environment for the Biologist. Proteins-Structure Function and Genetics 32(1), 1–2 (1998)
17. Rifaieh, R., Unwin, R., Carver, J., Miller, M.A.: SWAMI: integrating biological databases and analysis tools within user friendly environment. In: Cohen-Boulakia, S., Tannen, V. (eds.) DILS 2007. LNCS (LNBI), vol. 4544, pp. 48–58. Springer, Heidelberg (2007)

18. Tamura, K., Dudley, J., Nei, M., Kumar, S.: MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Molecular Biology and Evolution 24, 1596–1599 (2007)
19. Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J.: Basic local alignment search tool. J. Mol. Biol. 215, 403–410 (1990)
20. Kang, H., Lundeberg, M.A.: Participation in science practices while working in a multimedia case-based environment. Journal of Research in Science Teaching 47(9), 1116–1136 (2010)
21. Wolter, B., Lundeberg, M., Bergland, M., Klyczek, K., White, C., Tosado, R., Toro, A.: The influence of alternative instructional methods on student retention in postsecondary biology programs. Paper Presented at the American Educational Research Association, Denver, CO (2010)

User Interface and Information Management of Scenarios

Robert Louden, Matt Fontaine, Glenn A. Martin, Jason Daly, and Sae Schatz

Institute for Simulation & Training,
3100 Technology Parkway, OrlandoFL 32826
{rlouden,mfontain,martin,jdaly,sschatz}@ist.ucf.edu

Abstract. Scenario-based training has shown to be an effective instructional approach. However, it can be difficult for instructors to create scenarios as they are time-consuming to develop. This can result in the same scenarios being re-used, reducing training effectiveness. Furthermore, scenarios are based on training objectives and can vary in complexity. Researchers have investigated methods to simplify this process by assessing various approaches to scenario generation. Within such a tool, however, there needs to be a way to easily understand the different scenario components and their relationships within the scenario. This paper discusses the PYTHAGORAS (Procedural Yielding Techniques and Heuristics for Automated Generation of Objects with Related and Analogous Scenarios) system and specifically its editor used to create scenarios. We designed this system to simplify the process of scenario generation, providing a clearer understanding of the scenario components and relationships.

Keywords: Scenario Generation, Scenario Editor, Scenario Components.

1 Introduction

There are many aspects that make up any one training scenario. These can include environmental parameters, actions needing to take place, objectives to be accomplished, and many other factors based on the specific training. These components are indeed essential to create a variety of training exercises, which is important in providing improved training effectiveness. However, scenario generation is quite costly, which currently results in very small, haphazard scenario libraries and, therefore, reduced training [1]. There is a need for a faster and more efficient way to create scenarios that are qualitatively similar (i.e. fulfills the training objectives) yet appear to the trainees as different. However, the instructor must also be able to easily construct and understand the scenario being developed.

We designed the PYTHAGORAS system to meet these expressed needs. This system uses a scenario component approach to create custom scenarios and contains a scenario editor in order to do so. Using the visual representation of a mind-map (a diagram used to represent information branching from a central idea), the system provides improved comprehension of the scenario, its components, and their relationships to each other. The scenario editor also provides an intuitive interface for manipulating the different components.

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 30–39, 2011.

2 Visual Representation of Data

We have developed PYTHAGORAS to address the problem of creating scenarios that are qualitatively similar, in a quick and efficient manner. However, setting up a scenario with all of its different features can still be confusing. As the scenario becomes larger and more complex, it becomes more difficult for the instructor (e.g. scenario creator) to understand the different components and their relationships within the scenario. There needs to be a more effective and clearer way to present all of these data to the user.

The visual representation of data aids in the communication of information through illustrated means. The illustrations can be presented in a number of ways, including: diagrams, videos, and charts. These visual aids can allow for a clearer understanding of the subject matter. Visual representations of data are effective ways to both analyze and interpret data. The visualization highlights trends and patterns in numeric datasets that might not otherwise be immediately apparent [2]. Using a visual presentation, one could neatly display the information within a scenario. Yet which one would best display this information? As stated, there are several forms of data representations, so finding the most effective one can be difficult.

2.1 Scenario Editor

The training scenario starts with an initial situation, which represents the initial environmental setting of the overall scenario [3]. It includes the virtual environment being used and includes such parameters as time-of-day and wind speed. During the course of the exercise, additional information is added to the scenario in order to increase the complexity of the training and ultimately meet the objectives of the exercise. The additional information can include new environmental conditions, entity behavior, and events to take place. This various new information affecting various parts of the scenario can lead to a large collection of data. There needs to be a way to neatly display the scenario information and at the same time make it so the user can easily edit it if needed.

PYTHAGORAS allows for the creation of scenarios through the scenario editor, which gives a graphical representation of the overall scenario. This gives the data an aesthetic look and allows for the designer comprehension of what the particular scenario contains.

The initial version of the scenario editor (Fig.1) was time-based. The width of the graph signified the length of time for which the given scenario was to take place. Each event was stacked on top of the others, with the base situation (e.g. virtual world setting) at the bottom, which extended the entire width of the graph. The events were displayed as boxes with their size indicating its length in time. However, this approach provided some limitations to the system and editor. One of the disadvantages to this approach was that the scenario events do not always take a consistent time; there may be a minimum time at which the trainee must complete the scenario but never an exact time at which that event may end. Scenarios can also range down different paths, being that there are multiple ways to complete an objective, and some events may have prerequisites that must be completed or achieved before they can arise.

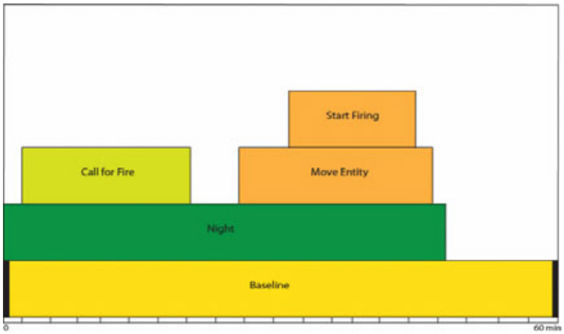


Fig.1. Time-based scenario editor. Figure shows that the scenario takes place for a total of 60 minutes (*bottom of graph*).

Based on the confusion from the disadvantages of the initial version, the editor needed to be improved upon, aiming to solve its issues and constraints. It needed to be user-friendly, not time dependent, better represent the scenario events, show that events can be dependent on other events, represent branching for the multiple paths a scenario can take, and have simpler and easy-to-use controls. The rest of this section discusses the different ideas for visualizing scenario components within the scenario editor, the advantages and disadvantages of each and why the resulting display was eventually chosen.

2.2 Web of Dependencies

The graph featured in Fig.2 displays the events connected to each other, based on their order in the scenario, resulting in a ring. Lines would spring from the events, pointing to the other events that were dependent on it. The graph clearly displayed the different dependencies of each event as a recognizable chord in the circle. The circle would then rotate based on the current event that was being edited, which would be displayed on the north cardinal point of the circle. It also gives a general concept of time, lining up the events one after the other based on order, yet without defining the time. However, despite these advantages, we decided that the graph was too confusing and did not show the possibility of events branching off to different paths.

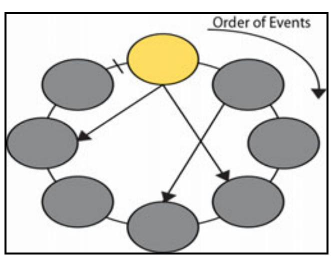


Fig. 2. Shows dependent scenario components (*nodes that are being pointed to*) and initial situation (*topmost node*)

2.3 Multi-dimensional Plot Graph

Using the idea of a plot graph, this approach consists of a multitude of graphs that can show the different details of the scenario by making them dimensions of their respective graph. This would enable the dimensions to be interchangeable so that it may show the scenario in different ways. Example graphs included difficulty versus dependency (Fig.3), where the fork of an event showed which event was more difficult than the other. The graph shows a good comparison between the two fields of interest. However, this approach is also confusing due to the excessive number of graphs.

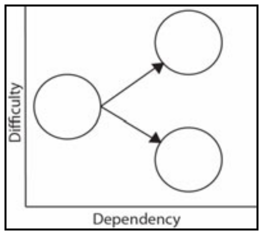


Fig. 3. Scenario editor as plot graph. This is an example of one of the details made for the scenario.

2.4 Assembly Line Plot Graph

The idea of this graph was to show the order in which events happen but also, which ones happen at the same time. With the width of the graph showing the time in a scenario, each event would spawn underneath the line at the particular time it would take place. The graph, shown in Fig.4, created presented a neat arrangement with a good amount of room for new items and displayed a clear flow of events, as well as time. However, this approach did not show event dependency or branching off to multiple paths and we found it to be increasingly difficult to understand as scenarios grew.

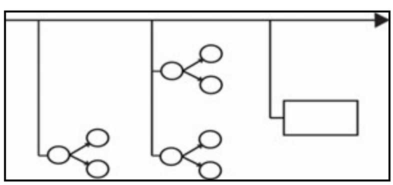


Fig. 4. Based on original editor showing time (*from left to right*) of scenario

2.5 Mind Mapping

This graph shows and represents the events around a central idea, or in this case the initial situation of the scenario. In order to avoid confusion and not overwhelm the user with large scenarios, the graphs can be simplified by showing and hiding the details of parts of the map. The overall scenario can be displayed as a whole or divided into components to show expanded detail.

The mind map has many advantages while avoiding the disadvantages of the others. Mind maps are an effective tool for planning and managing a particular problem or situation with many different aspects to it. It helps the user to rapidly build an enhanced understanding of any problem, challenge, or situation [4] [5]. It appears to be the best representation, in which we have found, of the scenario as it allows the user to plan and develop the different aspects of the scenario with great effect and efficiency, much like they would for any other problem. Certain challenges that come when developing a scenario include the many changes made to it and its complexity. Through the basic nature of a mind map, the scenario editor is able to address these challenges. In the end, we chose to use the mind-map display for our scenario editor.

The scenario editor displays the mind map in a tree like fashion, with one root preceding all other nodes in the tree. The tree is set up from left to right, the root starting at the left, rather than the conventional top to bottom (Fig.5 & 6). The root of the graph represents the initial situation and each node spawning from the root represents a particular component that makes up a scenario.

This new form of the editor led to the creation of newly defined components, in order to help expand the understanding of the different details that comprises a scenario. Further details about these components are stated in the next section.

3 Facets

PYTHAGORAS defines *facets* as the different components that make up the training scenario. There are currently a total of five facets: *baseline, augmentation, trigger, adaptation,* and *vignette*; all representing a particular characteristic of the overall scenario. Martin et al has a detailed review of all the scenario facets [6].

3.1 Baseline

This represents the initial situation of environmental elements of a scenario. The initial situation is the environmental setting supporting the elements required by the selected training objectives. Such elements include the time of day, weather, and location. There can only be one baseline in a scenario; thus it is at the root of the graph. Baselines, while basic, can support training. However, they are very simple scenarios by themselves. Additional elements need to be added in order to increase the complexity and fulfill additional training needs.

3.2 Augmentation

This facet is used to enhance the baseline scenario. The enhancements refer to alterations to the scenario by manipulating either the environment or entities. Examples of this facet include changing day to night, adding an additional target, and having high winds for the weather. In addition, some of these facets are placed with limitations of the number that can be added in any one scenario (e.g. It makes no sense to add "night" more than once). It should also be noted that augmentations focus on adding to the initial situation of the scenario.

3.3 Trigger

These facets are simply checks to determine if a particular event has taken place. This facet can either be event-based or time-based, determining if a specific event has happened or allowing for a time-specific event to happen. It also signifies a prerequisite for an adaptation. When the trigger is found to be true, the adaptations connected to it are then administered and any other triggers that may be connected to it as well will then begin to watch for their own specified event.

3.4 Adaptation

These facets are alterations to the scenario that will take place during the exercise, once the event illustrated in the trigger has occurred. They can include entity manipulations (create, kill, move, fire weapon) and environmental manipulations (reduce rain, raise sun). They can be used to adjust the elements of the scenario (possibly to adjust the complexity or to simply cause an event to occur) or to repair an exercise (in the case when an important entity was killed and would cause the exercise to otherwise end).

3.5 Vignette

The vignette is a combination of triggers, adaptations, and/or augmentations. It adds content tied to learning objectives to the baseline. Vignettes are defined as sets of the other non-baseline facets (e.g. augmentations, triggers, and adaptations) and can be both microadaptive (i.e., predesigned to meet instructional needs of trainees) and macroadaptive (i.e., adjusting in real-time to better facilitate the training). Each vignette contains an additional level of complexity and contributes to the total complexity of the scenario [7].

3.6 Facet Relationships

The scenario editor we created helps show the relationship between the different facets, connecting one facet to the next, as in a mind map graph. A line is connected between two facets, signifying a relationship between the two. The relationship itself varies in meaning based on which facets are connected. An example includes the relationship between triggers and adaptations. As stated before, for an adaptation to happen, a trigger must first be invoked. This understanding is seen graphically as the adaptations are connected underneath the appropriate trigger. More than one adaptation can appear underneath a trigger showing that multiple actions will take place once the trigger is invoked. However, only one of the adaptations is connected directly to the trigger. The others are connected to each other, one after the other in a line, in order to simplify the graph, and keeping a clean look.

In much the same way triggers and adaptations are connected, baselines and augmentations share the same appearance. Augmentations can only be attached to baselines as their role is to enhance the baseline scenario. In the exact same way adaptations are connected to triggers, multiple augmentations can be attached to the baseline, showing the multiple enhancements to the scenario. Since Vignettes are simply the combination of the other non-baseline facets (triggers, adaptations, and augmentations), they are attached only to the baseline and other vignettes. The relationships between vignettes show which vignette will be fired off next when the preceding one has taken place.

3.7 Multiple Editors

There are two forms of the scenario editor that PYTHAGORAS contains: one for creating a vignette and the other for creating the scenario itself. The typical user will use only the latter to develop scenarios based upon the current library of vignettes. However, an advanced user may use the former to create new vignettes to add to the vignette library. In the main scenario editor, the only two facets that are available to the user are baselines and vignettes. Similarly, in the vignette authoring editor, only triggers, adaptations, and augmentations are used. Both editors look and work the very same way so as to keep ease-of-use.

Aside from the facets available, the other main difference between the two editors is the root of the graph. The scenario editor has the baseline as its root whereas the vignette authoring editor simply has a placeholder that connects the three facets together. Triggers are connected to the right of the baseline placeholder whereas augmentations appear directly beneath it. Fig.5. shows an example of the main scenario editor and Fig.6 shows an example of the vignette authoring editor.

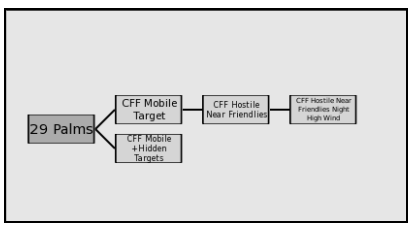


Fig. 5. This figure shows the scenario editor. The baseline facet (*root facet to the left*) is connected to the vignettes (*to the right)*.

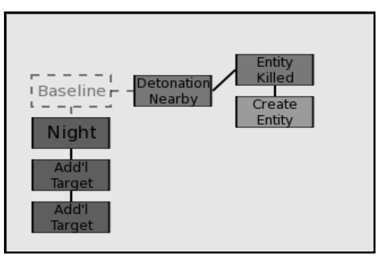


Fig. 6. This figure shows the vignette authoring editor. The baseline root (*dotted box)* is merely a placeholder to connect all the other facets and show their relationship between each other.

4 User Interface

The scenario editor provides an intuitive way for the user to create and manage the scenario and its facets. As already mentioned, the scenario editor has a very aesthetic look to it so that it can be easily comprehensible for the user to know all of the different aspects of the scenario. The graph portion of the editor itself takes up the majority of the screen space in the program (Fig. 7). To the left of the graph is a toolbar which lists the complexity level of the scenario that the user can select, selectable objectives for the scenario being created, and the list of facets that can be added. To the right of the graph is another vertical toolbar that displays the list of requirements that need to be fulfilled, based on the facets added. The requirements are additional information that needs to be defined, based on the training objective and added facets. It is worth mentioning that the list of facets available is based on the scenario's objective(s). Each facet contains information on which training objective it "supports" (meaning in which objective the facet is usable). The training objectives are enumerated based on the trainee audience and provided for user selection. [8]

The editor automatically adjusts itself to keep up with the aesthetic look, anytime a change is made to the graph (adding, removing, or moving facets). Each facet has a certain bulk that surrounds the box node to make sure that no other facet or lines overlap with each other.

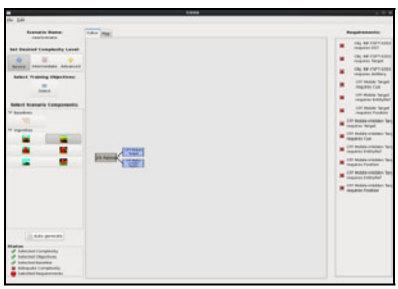


Fig. 7. Example scenario generation system built on top of PYTHAGORAS

4.1 Adding Facets

Every facet is represented by a node in the graph which has exactly one parent and one or more children to which it is connected with the exception being the baseline having no parent (since it is the root). For the scenario editor, the only facets available are the baselines and vignettes. So when adding facets to the scenario, the user must first select a baseline from the toolbar. The baseline is automatically added to the graph once selected. There can only be one baseline for a scenario; therefore, no others can be selected once one has been added, unless the current baseline is removed. Once a baseline is added, the user can select any of the available vignettes. To add a vignette the user simply selects one of the available ones from the list and then clicks on the facet in the graph to which it should be attached.

The vignette authoring editor works exactly the same way. Instead of baselines and vignettes, the facets available are augmentations, triggers, and adaptations. Augmentations are automatically added once selected, like baselines, and triggers work in the same fashion as vignettes. However since there is no baseline for this editor (just a placeholder) the first trigger selected is automatically added to the graph and acts as the root for all other triggers. Adding additional triggers works the same as adding vignettes in the scenario editor. To add an adaptation, the user selects one from the list and clicks on the desired trigger to attach it.

4.2 Removing Facets

PYTHAGORAS makes use of the right mouse button for removing facets from the graph. The user simply has to right click on the facet they wish to remove, which will bring up a menu of options. There is currently only one option listed in the menu, which is to remove the selected facet. Once removed, the graph will automatically adjust itself to compensate for the removed facet. The system also has a selection tool, allowing the user to select more than one facet at the same time. This in turn allows the user to remove more than one facet at once if desired. The selection tool works in a manner similar to most modern file browsing systems.

4.3 Moving Facets

The editor also allows for facets to be moved around the graph. For example, if an adaptation needs to be moved from one trigger to another the user can remove it and then add it again to the new trigger or the user can simply drag the adaptation from one to the other. The graph will add the facet to its new parent and automatically readjust the graph. This also works with the selection tool, moving more than one facet at once. However, the system does check to make sure that only valid facets are being moved. For example, if an augmentation, trigger, and baseline are selected and moved, the system will know that a baseline cannot be moved (since it is the root) and that an augmentation and trigger cannot be moved at the same time.

5 Conclusion

The more information required for a scenario, the more difficult it can become to create and understand everything in it. Data visualization aids in the process of being able to easily comprehend the overall scenario. Different forms of visualization work well for different forms of data [2]. A mind map approach appears to work well for representing scenario components graphically [5]. PYTHAGORAS uses this form to aesthetically display the scenario's information through the use of its scenario editor. The editor itself aids in the process of the scenario's creation, reducing the time it normally takes to create one, while graphically showing what the scenario is made up of and how all of its components are related to each other and the scenario.

We now plan to evaluate the developed mind map visualization approach to scenario representation. While we feel the approach taken fulfills the needs of scenario generation, an experiment is planned to test the efficacy of our scenario generation approach (including the data visualization component). In addition, we wish to explore mind map representation for handling more complex groupings of scenario facets (e.g. our current approach does not handle well the case of two triggers that must occur but where order is not important).

References

1. Chipman, S.F.: Putting the "advanced" distributed learning. In: O'Neal, H.F., Perez, R. (eds.) Web-based learning, pp. 143–162. Lawrence Erlbaum, Mahwah (2006)
2. Egger, A.E., Carpi, A.: Data: Using Graphs and Visual Data. Visionlearning POS-1(4) (2008)
3. Tomizawa, H., Gonzalez, A.: Automated Scenario Generation System in a Simulation. In: Proceedings of Interservice/Industry Training, Simulation and Education Conference, National Defense Industrial Association, Orlando
4. Farrand, P., Hussain, F., Hennessy, E.: The efficacy of the 'mind map' study technique. Medical Education 36, 426–431 (2002)
5. Tucker, J., Armstrong, G., Massad, V.: Exploring the World of Mind Maps
6. Martin, G., Schatz, S., Hughes, C.E., Nicholson, D.: What is a Scenario? Operationalizing Training Scenarios for Automatic Generation (July 2010)
7. Martin, G., Schatz, S., Bowers, C.A., Hughes, C.E., Fowlkes, J., Nicholson, D.: Automatic Scenario Generation Through Procedural Modeling for Scenario-Based Training. In: Proceedings of the 53rd Annual Conference of the Human Factors and Ergonomics Society (2009)
8. Martin, G., Hughes, C.E.: A Scenario Generation Framework for Automating Instructional Support in Scenario-based Training (April 2010)

Giving UI Developers the Power of UI Design Patterns

Jocelyn Richard[1,*], Jean-Marc Robert[1], Sébastien Malo[2], and Joël Migneault[2]

[1] École Polytechnique de Montréal, P.O. 6079, St. Centre-Ville
Montréal QC, H3T 1J4, Canada
{jocelyn.richard,jean-marc.robert}@polymtl.ca
[2] CAE Montréal, 8585, Chemin de la Cote-de-Liesse
Saint-Laurent, QC H4T 1G5, Canada
{malo,joelmig}@cae.com

Abstract. UI developers interested in UI design patterns generally face major problems when trying to use them, because of the complexity of pattern libraries and the lack of supporting tools. As a consequence, UI design patterns are not widely used and this entails an important loss of productivity and quality. In this study, we identified and wrote 30 UI patterns that were made available in a library, and we compared four modes of presentation for them: pattern thumbnails, application types, decision trees, and alphabetical mode. Ten subjects participated in the study. User satisfaction was higher with the three new modes than with the alphabetical mode. Search time was higher with the three new modes than with the alphabetical mode. Although difficult to evaluate, pattern relevance was better with the three new modes. Those findings were turned into recommendations for immediate applications. In the conclusion, we propose some research avenues for the future.

Keywords: user interface design patterns, design pattern libraries, pattern language user interface design, mode of presentation.

1 Introduction

« Each pattern describes a problem which occurs over and over again in our environment, and then describes the core of the solution to that problem, in such a way that you can use this solution a million times over, without ever doing it the same way twice. » [1]

Design patterns capture invariants of a proven *solution* to a recurring *problem* in a specific *context*. They are reusable knowledge blocks, written in plain text following a predefined format, and possibly illustrated by photos, code samples, diagrams, etc. Patterns are often organized in a *language* that is related semantically and hierarchically, in order to generate complete design solutions [1].

Christopher Alexander set a landmark in architecture when he wrote the first patterns, 30 years ago. His idea has since spread into the fields of pedagogy, business processes, object-oriented programming and Human Computer Interaction (HCI).

---

[*] Mr. Richard is doing a Master's degree in Industrial Engineering at Polytechnic and a 1-year internship at CAE.

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 40–47, 2011.

The use of user interface (UI) patterns remains low, in spite of the large quantity of available pattern languages [2]. Indeed, interface developers do not use UI patterns much for various reasons:

- A *lack of organization*: as a pattern language grows, it can accommodate more than a hundred patterns. Without clear entry points, shortcuts, or guidance, finding the right pattern often amounts to skimming through the entire list. This heavy work load adds to the regular tasks of developers.
- The convolution *of current patterns*, written by « usability engineers or experienced UI developers » [3] for themselves [4]. Interface developers are not familiar with the concepts, methods and vocabulary used in human factors engineering; thus it is difficult for them to estimate the relevance or applicability of a pattern to a situation [3]. They also have difficulties understanding how patterns are related and how to benefit from these relations.
- A *lack of tools* for supporting interface developers [5] and reducing their workload.

The difficulty of using UI patterns may entail a huge loss of productivity and quality for interface developers who actually code the UI, as they pass over a tool that can « improve the consistency of the user interface and UI quality, make designing and implementing more efficient, provide business advantages for software companies, and work as a common language communicating the design knowledge to the designers, developers, and other stakeholders » [6].

2 Our Approach to Write and Present UI Patterns

Our research work on UI patterns is based on four main activities: an extensive literature review which allowed us to draw lessons from others' work on UI patterns, an analysis of different training and simulation tools interfaces developed at CAE, discussions with human factors specialists who knew the needs of UI developers at CAE and were looking for more rigor and consistency in UI, and discussions with several UI developers at CAE since they are the end users of the UI patterns.

We wrote 30 UI design patterns for training and simulation support tools developed at CAE, and shared them on the company network on a need-to-know basis. They have been actively used and validated by more than 50 engineers, human factors specialists, and UI practitioners, and tweaked in the light of the feedback of these professionals.

We wrote UI patterns following the usual « Problem-Context-Solution » format, with a heavy use of bullet points (rather than dense paragraphs) and example imagery (up to half the content of the pattern). This aimed at increasing legibility, understandability and appeal of the patterns for developers.

Furthermore, UI patterns were structured according to three modes of presentation we developed: *application types*, *decision trees* and *pattern thumbnails*.

Application types are halfway between *postures* as defined by Welie [7] and *Standard Screen Patterns* as defined by Neil [8]. They are textual and visual representations of stereotypes or software genres like Web browser, spreadsheet, media player, to which a predefined pattern collection is linked. As the quantity of application types is inferior to the quantity of patterns, and as they are a more user-friendly artifact, they can act as a shortcut during pattern identification and search operations.

Decision trees replace the difficult processing of a pattern [9] with a succession of questions requiring less time and HCI expertise for finding a pattern. Besides, this mode makes explicit otherwise hidden pattern relationships. Decision trees look like step-by-step wizards (Figure 1).

Wizard	< Back	< Back
What is the function of the child window?	**What is the nature of the message?**	Use a Notification Stack.
Display more features > Such as toolbars, lists, menus...	Alert or warning > « Low Battery »	
Send the user a message > Warning before deletion, completion.	Information or confirmation > « Update complete »	

Fig. 1. A possible path for the « Replace child window » decision tree, showing what happens if the user clicks on « Send the user a message » on the 1st step and « Information or confirmation » on the 2nd step

Pattern thumbnails are loosely based on the idea that a pattern can be referenced to by a rich snippet of information [10]. They combine the pattern's title to its common denominations, a one-liner and a schematic display of the solution (Figure 2). This thumbnail provides a cognitive shortcut to the identification and understanding of the pattern by the developer.

Direct manipulation

Create, edit or browse objects by acting « physically » on them rather than using toolbars.

Fig. 2. Example of a pattern thumbnail for the « Direct Manipulation » pattern

The goal of our study is to compare the performance of four modes of presentation of UI patterns: the three modes described above and a reference mode, which is a one-page alphabetical index linked to all patterns in the database, much like in Welie's, Tidwell's or Malone and Crumlish's libraries[1].

3 Methodology

Subjects. Ten professionals from CAE participated in the study. Four are software developers, four have team or project management responsibilities, and two have graphical or quality assurance tasks. None of them had heard of UI design patterns or seen CAE's patterns library before. The age varies between 25 and 53 years old.

Task. Subjects were asked to find, with each of the four modes of presentation, a pattern suitable to a specific CAE user interface they had freely chosen at the beginning of the test. « Suitable » could either mean: « That would significantly improve the chosen UI » or « That is already and correctly applied in the chosen UI ». The subjects were not asked to technically apply the pattern, since this study only focuses on evaluating the capability of each mode to help one find rapidly a relevant design pattern. The library contained 30 UI patterns; each time the subject found a pattern with a mode, the same pattern could not be used with another mode and the different modes were used in a random order by the subjects.

Data. Data collected about each mode were about these parameters: the time required to find a suitable pattern, the user satisfaction about the mode, and the relevance of the selected pattern for the initially chosen CAE user interface. The data about the last two parameters were collected through an interview at the end of the session. We evaluated the subjects' first-contact experience with the library: if they could find a relevant pattern on their own and were interested in using it. The subjects were asked to comment aloud their activities while examining the pattern library and choosing a pattern.

Procedure. At the beginning of a test session, we welcomed the subject, explained the goal of the study, had the subject complete a short biographic questionnaire, and gave brief instructions about the task to perform and the sequence of events during the session. The test was held at the subject's desk, in the presence of the experimenter (the first author of this paper) who observed the work done by the subject, collected the questionnaire, and asked questions.

4 Results

Subjects were very successful for finding UI patterns. They found patterns 37 times out of 40 trials: the three failures are due to work-induced interruptions of the trial.

[1] www.welie.com/patterns/
http://designinginterfaces.com/
www.designinsocialinterfaces.com/patterns/

User satisfaction is higher with the three new modes of presentation than with the alphabetical index (see table 1 and table 2): with the application types and the decision trees it is 42% higher (3,7 vs 2,6), and with the thumbnails it is 73% higher (4,5 vs 2,6). Moreover, the standard deviations is lower for the ratings of the new modes, especially the thumbnails, than the rating of the alphabetical order.

On the other hand, the search time of the UI pattern is higher with the new modes (see table 3): 11% higher with thumbnails (03:26 vs 03:06), 23% higher with decision trees (03:48 vs 03:06), and 51% higher with application types (04:41 vs 03:06). We will see below that this may comprise some advantages. Standard deviations are similar this time, nearly one minute for every mode.

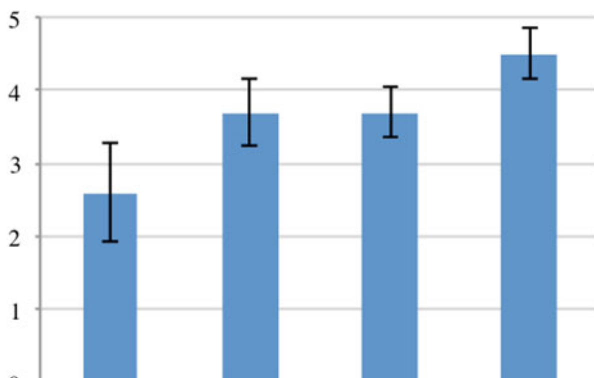
Thumbnails have the highest user ratings for satisfaction and are the favorite mode of 7 subjects out of 10 while being the second best mode for searching time. Conversely, alphabetical index is the most rapid mode but with the worst rating of user satisfaction.

Design patterns found with application types and decision trees modes were generally more relevant to the benchmark UI than patterns found with thumbnails and alphabetical index. It is worth reminding that all the modes were referring to the same pattern library. The relevance of a pattern for a given UI is somewhat subjective and

Table 1. Modes of presentation preferred by the subjects (each participant had to pick one out of four)

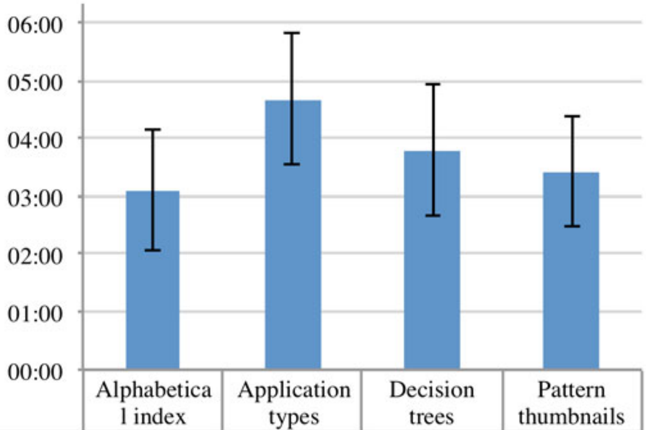
Alphabetical index	Application types	Decision trees	Pattern thumbnails
0	3	0	7

Table 2. User satisfaction for four methods (Likert scale from 1 to 5, the higher the better)



	Alphabetical index	Application types	Decision trees	Pattern thumbnails
Average	2,6	3,7	3,7	4,5
Standard Deviation	1,35	0,95	0,67	0,71

Table 3. Time required to find a pattern (the lower the better)



	Alphabetical index	Application types	Decision trees	Pattern thumbnails
Average	00:03:06	00:04:41	00:03:48	00:03:26
Standard Deviation	00:02:06	00:02:16	00:02:16	00:01:57

hard to agree upon, yet the findings show a clear trend: patterns found with thumbnails and alphabetical index methods were generic, including all-purposes recommendations such as Good typography or Good color. Whereas patterns found with application types and decision trees modes, like Cloak of shadows or Mega drop-down, were more specific and with a greater value for design.

In addition, most of the subjects have praised at length the choice, quality, and quantity of examples in the patterns as well as the bullet writing style even though these particular aspects had not been mentioned by the observer.

5 Discussion and Recommendations

5.1 Discussion

The overall response to the proposed approach, new modes of presentation and writing style, has been very positive.

CAE UI patterns library was deemed attractive, and the subjects, with no previous knowledge of UI design patterns, were able to find patterns relevant to their situation with a high satisfaction rate. The search time was higher with the new modes but this result actually seems to be a good thing: excess time wasn't spent on hit-or-miss, on trial-and-error blind exploration of patterns as with the alphabetical index, but rather on reading additional details or thinking about the problem. Context and guidance provided by the new modes of presentation may require more time, but this could be readily accepted by the subjects as it leads to more relevant patterns and a better understanding of UI issues.

Furthermore, application types and decision trees act as multiple and targeted access points to subsets of the library, *de facto* chunking it for UI developers. This chunking, along with the context provided through textual and visual documentation, contribute to building the subjects' mental model of the UI principles. It also decreases the workload of using the library: developers do not have to know anymore the library by heart to select the right pattern. Recognizing the problem they are confronted to or the type of application they are working on is enough to provide a solution.

The complementarity of the modes was also noted. Application types received better ratings from subjects involved in new products, where the UI had yet to be designed, because they gave them a starting point. On the contrary, the decision trees received better ratings from subjects involved in mature products with specific UI problems.

The pattern thumbnails were really appreciated and are useful on two levels: first as an introduction to the library, playing on the user's curiosity (subjects enjoyed guessing the meaning from the thumbnail and spent time doing so even when they already had found a suitable pattern); second, as a nice quick reference for advanced users, providing rich information at a glance.

Yet focusing solely on thumbnails would be a mistake. Application types and decision trees, although less ostentatious, provide more accurate answers and much more context, which is beneficial in the long-term to the UI knowledge of software developers.

Most subjects expressed their liking of several modes at the same time and urged us on not picking one at the expense of the others.

5.2 Recommendations

- Supporting each pattern presentation with a pattern thumbnail.

  It is a low-cost yet effective way of achieving findability and it conveys a lot of meaning. The visuals used should be schematized in order not to silently enforce a particular visual style nor a specific technical solution.
- Support the library with application types and decision trees.

  They are the modes yielding the best results, but since they are harder to define and write, it is better to first have a steadfast library and a fine understanding of the organization to build upon.
- Keep an alphabetical index available.

  Even if they do not use it, subjects feel safe having a « no surprise » solution available. Some users use it to gauge the scope of the library at a glance.
- Write patterns using bullet points and several carefully chosen examples.

  They make patterns easier to grasp for an audience that has no training in HCI.
- Keep resources for « the usability of usability ».

  The new browsing modes mean added writing, coding and drawing. This may constrain a team with limited resources to write fewer patterns *per se* in order to make the remainder accessible to interface developers.

6 Conclusion

Our results are encouraging, even though not all positive, since the subjects who participated to the study were able to find UI patterns on their own, their satisfaction with three new modes of presentation of patterns that we devised was higher than with the conventional alphabetical mode, and the relevance of patterns they found with these new modes was better than with the alphabetical mode. Hence *application types*, *decision trees*, *pattern thumbnails,* and a very visual writing style can be considered as effective and reliable to make patterns easily accessible to UI developers with no prior HCI training. The four modes of presentation we tested are not exhaustive and it is worth making the effort to devise and test new ones for the benefits of UI developers, and of other stakeholders o the enterprise (e.g., clients, partners).

Finally, the « language » aspect of design patterns has been deliberately left out of this paper, mainly because the accepted definition of pattern language varies widely. As of now, our patterns are indeed naturally structured around application types and decision trees that perform like thematic « hubs ». However, this way of working only accounts for a small part of the functions of a pattern language; here lies a huge area of improvement.

References

1. Alexander, C., Ishikawa, S., Silverstein, M.: A Pattern Language. Oxford University Press, Oxford Oxfordshire (1977)
2. Mahemoff, M., Johnston, L.: Usability pattern languages: The “Language” aspect. In: Human-computer interaction: INTERACT 2001: IFIP TC. 13 International Conference on Human-Computer Interaction, pp. 350–359 (2001)
3. Seffah, A., Javahery, H.: On the usability of usability patterns. In: Workshop entitled Patterns in Practice, CHI, pp. 1–2 (2002)
4. Tidwell, J.: Designing Interfaces: Patterns for Effective Interaction Design, p. xi. O’Reilly, Sebastopol (2005)
5. Breiner, et al.: PEICS: towards HCI patterns into engineering of interactive systems. In: Proceedings of the 1st International Workshop on Pattern-Driven Engineering of Interactive Computing Systems, pp. 1–3 (2010)
6. Lammi, J.: Developing a UI Design Patterns Library – A Case Study at eCraft. Helsinki University of Technology, p. 2 (2007)
7. Welie, M., Van der Veer, G.: Pattern Languages in Interaction Design: Structure and Organization. In: Proceedings of Interact 2003, pp. 527–535 (2003)
8. Neil, T.: 12 Standard Screen Patterns. Designing Web Interfaces (2009), `http://designingwebinterfaces.com/designing-web-interfaces-12-screen-patterns` (retrieved October 28, 2010)
9. Gaffar, A., Seffah, A., Poll, J.A.V.D.: HCI pattern semantics in XML: a pragmatic approach. In: Proceedings of the 2005 workshop on Human and social factors of software engineering, pp. 1–7 (2005)
10. Meszaros, G., Doble, J.: A pattern language for pattern writing. Pattern languages of program design 3, pp. 529–574. Addison-Wesley Longman Publishing Co., Inc., Reading (1997)

The Cultural Integration of Knowledge Management into Interactive Design

Javed Anjum Sheikh, Bob Fields, and Elke Duncker

Interaction Design Centre,
School of Engineering & Information Sciences,
Middlesex University, London, UK
{j.anjum,b.fields,e.duncker}@mdx.ac.uk

Abstract. This research is being conducted to address the integration of cultural factors in interactive information system to enhance the reach of knowledge management to culturally heterogeneous user. In this context, we introduced a knowledge management framework. This method will help researchers to understand how different cultures view similar concepts. The result and analysis we expected in this research is intended to guideline to design knowledge management based cross-cultural interface. The result will increase usability enhancement and interaction patterns in interactive design.

Keywords: Cultural Integration, Cultural Interface, Cultural Representation, Cultural Design, Knowledge Management, Usability, User Interaction.

1 Introduction

This research focuses on cultural knowledge management and managing cross-cultural differences. The research describes emerging issue of cultural categorization with the context of Knowledge Management. The cultural and human context is often overlooked by Information System (IS) and Knowledge Management (KM) practitioners. The main problem is to capture cultural knowledge and its presentation. The purpose of this paper is to discuss cultural knowledge with respect to information technology. The truly effective cultural knowledge management is not only translating text but the development of a tool to represent similar cultural concepts for global user.

In what follows, we review related works on culture and design in section two. In section three we describe our research methods. Section four presents the results of our study and the analysis thereof. Following a discussion of links to interaction design in section five, our conclusions and directions for future work are presented in the last section.

2 Background

Cross-cultural issues exist in every aspect of our lives. When two peoples get together than culture clash starts. Normally people accommodate, sublimate or ignore culture

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 48–57, 2011.

difference due to common ground, shared goals same interest. The difficulties arise when people consider language, ethnic or racial background, religious beliefs, gender, age, education in depth. There are many avenues to take that will help cultures 'get along' better: identifying and concentrating on improved communication, fostering mutual respect, engendering co-operation and shared visions. Then, all kinds of cultures can work side by side to everyone's benefit. The concept of culture is most commonly associated with anthropology. Further disciplines using culture as a concept and thereby focusing on specific aspects of culture are sociology, psychology, economy, and computer science. Every person gets more satisfaction in their own cultural environment as it helps to increase efficiency and productivity. The studies by [1][2][3][4][5][6] show that culture plays a vital role to understand the psychology of any nation. The main problems area is cultural factors and their acceptance and without considering these cultural factors, knowledge management cannot be useful for everyone. The research on cultural differences [7][8], is inadequate in its context, whereas research investigating user point of view in cultural differences [9][10] needs attention.

All of these studies attributed differences to various cultural styles: including searching style, cognitive style, language use, perceptions of search systems and information sharing. The most important cultural clash is lack of culturally specific classification system. Classification is the process in which objects are grouped and classified which are rooted in people's experience. Classification systems are way of organising knowledge. Currently used classification schemes are Dewey Decimal Classification (DDC), Library of Congress Classification (LC), Universal Decimal classification (UDC), Cutter's Expansive Classification, (CEC), Bliss's Bibliographic classification (BBC), and Ranganathan's Colon Classification (RCC). They claim to be universal schemes but these suffer from an inherent bias. They have great possibilities of expansion, alteration and adjustment only for new western subjects. However these schemes are failed to classify local material in appropriate ways. This means that classifications made by one culture can be difficult to access for another culture Fig 1.

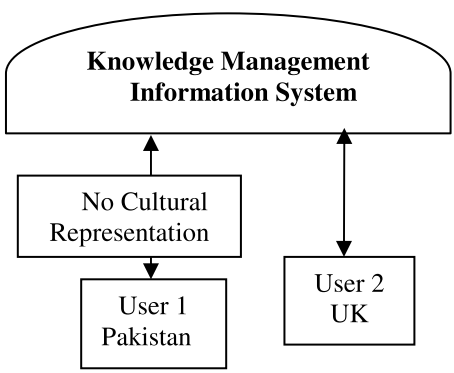


Fig. 1. The Problem

3 Methods

Researchers use Hofsted's Model without keeping in mind that his model is not only too rigid [15] but also has very general views [1] about different cultures. It is also not appropriate for Cultural based knowledge management system [16] being a business model [2]. Therefore, Hofstade's concept of cultural dimensions (Individualism vs. Collectivism, large vs small Power Distance, strong vs. weak Uncertainty Avoidance and Masculinity vs.Femininity) has limited impact for this research.

Cultural specific study is an important phenomenon in research, but computer scientists neither consider it to be an essential element for them nor do they think the user is important, [17][18].The results of studies by [13][19][6] show computer scientists face difficulties in the successful integration of culture into computer application

The knowledge management researchers focus their study on organizational culture [20][21][22][23].The researchers focus on explicit knowledge and failed to fully incorporate Nonaka and Takeuchi's model [24]. As we know there are two types of knowledge, namely tacit and explicit knowledge. Tacit knowledge resides in the mind of individual and is difficult to capture and document. Whereas Explicit knowledge can be captured and documented. We are using both types of knowledge to achieve successful cultural knowledge. The researchers focus on explicit knowledge and failed to fully incorporate Nonaka and Takeuchi's (2005) model. We are using Nonaka model to capture not only explicit but also tacit knowledge fig 2.

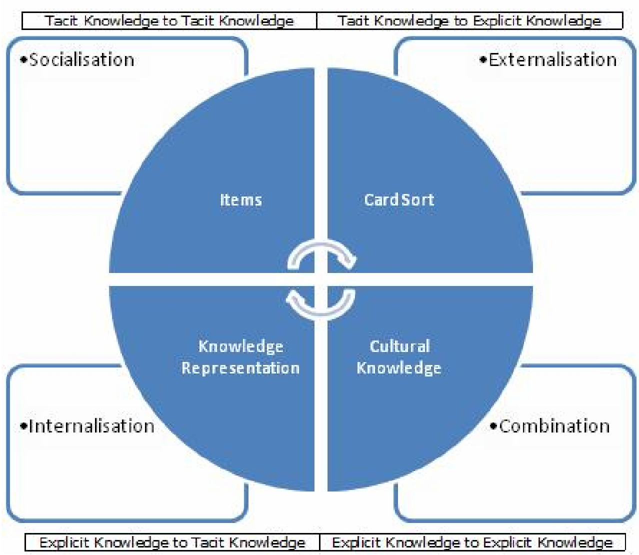


Fig. 2. Cultural Knowledge Management Model

The most important cultural clash is lack of culturally specific system. For instance, none of these systems are fully capable of classifying and organising material for cross-cultural user. In Western culture this problem is not very much acute but in non-Western culture, it is very confusing and perplexing resulted retrieval of information is not up to mark. Duncker, [25] emphasises on to create system for cross-cultural user. Therefore it is need to integrate cultural differences with interaction capabilities [26].

3.1 The Study

In order to approach the question of how different cultures organise their knowledge differently, we employ a research approach based on card sorting. In a card sorting study participants are typically asked to arrange into groups, cards on which the names of objects are written. Such studies can reveal something about the ways that participants organise their knowledge of the world. The approach is widely used for initial exploration, in the field of knowledge acquisition [27]. It helps to develop and identify concepts, models, attitudes, trends, patterns and values for capturing information from the mental model of the participants. The mental model generates concepts and suggests possible taxonomies.

Card sorting is widely used in Human Computer Interaction, psychology, and knowledge engineering for knowledge elicitation. It helps to evoke participants' domain knowledge [13] distinguish the level of the problem [28] and reflects ideas about knowledge [29]. This research investigates the cultural classification by using the Card Sorting technique as an effective method to identify categories of food items. Card Sorting is often used to gather data about personal construct for instance menu structure specification and to understand users' perceptions of relationships among items. These experiments use one layer of grouped card. Our card sorting technique is different. In that we in ask subjects to further classify the groups of cards, if they wish to do so.

The following method is adopted to conduct the studies. Thirty-nine printed cards were used, each carrying the name of a distinct food item that participants would be asked to group together with other items. The purpose of this study was to examine and demonstrate the differences in the way participants from Pakistan and the UK categorise entities. We chose Pakistan and the UK for field study, as both countries are multi-cultural, multi faith and have sub-cultures.

The **domain** chosen for this research is food items as these are very common to everyone. As the study was conducted in Pakistan, the food items were translated into the Urdu language as closely as possible.

3.2 Participants

A group of 160 participants (80 PK and 80 UK) took part in this research. The participants recruiting process in Pakistan ran smoothly, but we encountered difficulties in recruiting UK participants, due to criteria (of pure British). Because of the recruitment method, it was not possible to achieve gender-balance especially in the UK. None of the participants recruited for the Pilot Study was used again in the main study. No post study interviews were conducted

All participants provided written informed consent. The study was reviewed and approved by the Middlesex University Ethics Committee.

As the study was a cross-cultural one, the food items were translated into the participant's first language where appropriate. Participants were asked to classify items in several steps Figure 3. First they grouped items and gave labels to each group. Then they were asked to further divide the group if they wish to do so. The steps of grouping and labelling are repeated until the subject no longer wants to subdivide the groups of items. When subgroups were produced, the participant was asked to provide a name for that group. This usage is in contrast to many popular applications of card sorting, where participants are asked to organise the cards into groups, without further sub-categorisation to generate a deeper tree of classifications.

Fig. 3. Card Sorting experiment

4 Findings

4.1 Observations

This study shows a unique result compared to other studies. There is a large difference in categorization in the respect of culture, faith and environment. The following presents some examples from the studies to support this claim:

- Educated participants have western perception about prohibited items in Islam.
- Pakistani women are less privileged; they receive only religious education or basic education. Consequently, their perception of prohibited items is faith oriented.
- Participants faced difficulty to categorize items like Doughnut, Pasta, Pizza, as they are not familiar with these items.
- Few participants from Siraiki culture categorized items according to their profession
- It has been observed that this culture has a strong Islamic influence. The participants in this study have less western influence on their categorisation.
- Although some food items were categorized easily, some other items can be placed in two categories and some were unfamiliar to participants.
- Some items were ambiguous for participants and they tried to place them into two categories.

- Some participants were not familiar with some items, and interpreted them differently.
- Some participants sorted items into groups and sub groups whereas others in a more general way.

The initial analysis of the study indicates a cultural difference in food categorisation among people belonging to different cultures that appears to be greater than the differences between people within the same culture. The studies suggest that both the ‘national culture’ and the ‘belief system’ of a participant shape the way they categorise items. By ‘belief system’ here, we refer roughly to religious background as this is a highly significant factor in the way people understand food and the various domestic practices that surround it. It seems likely that other elements of culture, such as professional cultures or membership of communities of practice would gain greater significance.

5 Discussion

Based on our findings, we can more confidently assert that categorisation is influenced by culture and belief, and several open questions exist as to the nature of this influence.

A first step towards conducting analysis in a more rigorous manner was to formalize the complimentary notions of ‘similarity’ and ‘difference’ that are at work. A number of possible formulations are possible, but the one that proved to be most promising was the notion of ‘edit distance’: the difference or distance between two tree structured hierarchies is considered to be the number of editing steps necessary to transform one tree into the other. This measurement of distance was implemented in software based on a freely available framework called SimPack.[1]

The algorithm for computing the ‘edit distance’ between trees facilitated the construction of a ‘distance matrix’ that encodes the edit distance between the hierarchies produced between all pairs of study subjects, and the discovery of structure in the population of subjects entails an exploration of this distance matrix.

Two approaches to this exploratory task were employed. A more traditional statistically-based approach was implemented using a variant of the k-means cluster analysis algorithm to discover clusters of subjects who were ‘close together’ in that they produced similar hierarchies. This formal style of analysis was complemented with a more exploratory tool that produces a visualization of the distance matrix, based around the physical analogy of data points joined by a collection of springs whose length is determined by the edit distances[2]. A simulation of such a system yields a dynamic network that tends to settle in a ‘low energy’ configuration. The latter technique provides a useful visual way of seeing how a structure emerges from the confusion, as similarly similar trees tend to gravitate towards one another.

[1] SimPack is an open source collection of software tools for investigating the similarity between ‘ontologies’. Available from http://www.ifi.uzh.ch/ddis/simpack.html

[2] The tool is based on the Graph demonstration program that is part of the Java Software Development Kit available fromjhttp://java.sun.com

The results obtained from the cluster analysis algorithm are, informally, at least, in accord with the graphical simulation. Figure 5 illustrates this by showing the graphical display in which subjects are shown as numbered nodes in the graph. The physical distance between nodes in the figure is a reflection of the network of edit distance relationships. Overlaid on the figure are the four clusters found by a run of the cluster analysis algorithm.

A starting assumption for this research was that no scheme for organising information is likely to be equally effective for a range of cultural groups. The current research aims to make a contribution in this area, not by finding a universal way of classifying information, but by providing a method for investigating classification in a locale in order to generate localised interface designs. The expected solution will be based on local user access needs and capability of the local users.

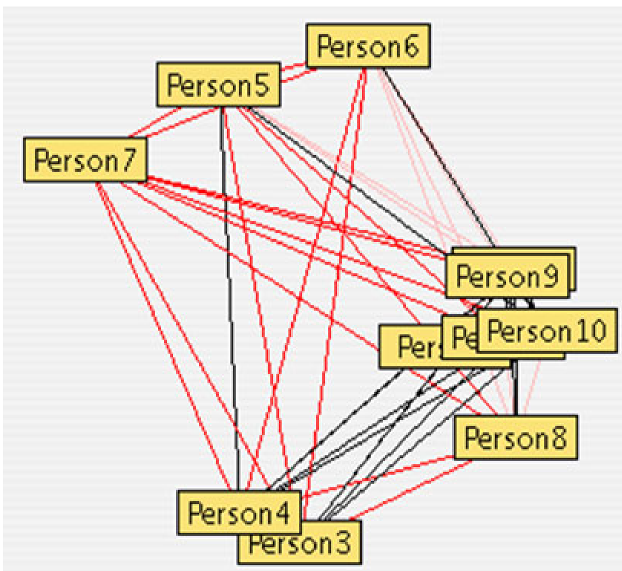


Fig. 4. Overlaying cluster analysis on a graphical representation

The analysis so far has given a way of identifying clusters of related structurings of a set of objects. A number of strategies exist to take this forward into interaction design. An obvious approach is to provide a localized user interface for each cluster, choosing a representative element from a cluster (for example, using the edit distance metric to find the most central element in a cluster) to guide the structure of navigation elements on an interface.

If they try to go beyond this point, they are automatically asked, whether they would like to order the pdf, and are given instructions as to how to do so.

Please note that, if your email address is given in your paper, it will also be included in the meta data of the online version.

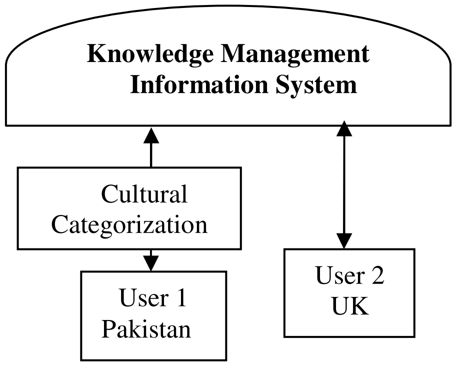


Fig. 5. The Solution

6 Conclusion and Future Work

The contribution of this research is to introduce a new analysis technique based on SimPack's modified classes to discover/measure similarity and edit distance. This method will help to understand how different cultures view similar concepts.

We propose a cultural based interface obtained from cultural knowledge. Our interface will show a common concept which is result of hierarchical clustering analysis of multi cultural representation. It will allow user to explore effectively in comparison to a non-cultural based interface.

The interface is user perspective, which will help the user to interact effectively and close to human to human interaction. When the users visit the main page the interface user will find cultural based navigations/classifications close to the particular culture. If the user clicks on a particular culture for example Pakistan, the user will get additional options from four cities. We hope this will enable the reader to understand the problem and its solution will give them better overview of this research. The aim of this research is to propose a design for all cultures to increase usability enhancement and interaction patterns in categorizing that lead to browser design. The result and analysis we presented in this paper is intended to guide design of cross-cultural interface. We used mixed methods to interpret the result. Our studies explored cultural difference by card sorting and result analysis through cluster analysis to compare both cultures. Significant differences were found in term of categorisation. The result increase usability enhancement and interaction patterns in categorising that lead to interface design.

References

1. Bourges-Waldegg, P., Scrivener, S.: Representations, the central issue in cross-cultural interface design. Interacting with Computers 9(3), 287–309 (1998)
2. Del Galdo, E., Nielsen, J.: International User Interfaces. Wiley Publishing, New York (1996)

3. Evers, V.: Human Computer Interfaces: Designing for Culture, Masters Thesis Faculty of Mathematics, Informatics, Physics and Astronomy, University of Amsterdam, Amsterdam (1997)
4. Minocha, S., French, T., Dawson, L.: Cultural Attractors of Usability and Customer Relationship Management in (B2C) E-Commerce Environments. In: Gunter, K., et al. (eds.) Culture and HCI: Bridging the Cultural and Digital Divides. Conference (6th: 2003). Proceedings of the 2nd British Computer Society Human Computer Interaction Workshop, pp. 37–47. University of Greenwich, London (2003)
5. Sun, H.: Building a culturally-competent corporate web site: An exploratory study ofcultural markers in multilingual web design. In: ACM Special Interest Group for Design of Communications. Conference (10th. 2001). Proceedings of the 19th Annual International Conference on Computer Documentation, pp. 95–102. ACM Press, New York (2001)
6. Yeo, W.: Global-software development lifecycle: An exploratory study. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 104–111. ACM Press, New York (2001)
7. Hofstede, G.: Culture's Consequences, 2nd edn. Sage, Beverly Hills (2001)
8. Masuda, T., Nisbett, R.E.: Attending holistically versus analytically: comparing the context sensitivity of Japanese and Americans. Journal of Personality and Social Psychology 81, 922–934 (2001)
9. Ein-Dor, P., Segev, E.: End-User Computing: A Cross-Cultural Survey. In: Proceedings of the Twenty-Third Annual Hawaii International Conference on Systems Science, vol. IV, pp. 240–250 (January 1990)
10. Igbaria, M., Zviran, M.: End-user effectiveness: a cross-cultural examination. Omega, Int. J. of Mgmt Sci. 19(5), 369–379 (1991)
11. Sheikh, J.A., Fields, B., Duncker, E.: Cultural Representation for Interactive Information system. In: Proceedings of the 2009 International Conference on the Current Trends in Information Technology, Dubai (2009)
12. Luong, T., Lok, J., Lok, S., Driscoll: Internationalization: Developing Software for Global Markets. Wiley, New York (1995)
13. Russo, P., Boor, S.: How fluent is your interface? Designing for international users. In: Conference Proceedings of the Conference on Human Factors in Computing Systems, pp. 342–347. Addison-Wesley Longman Publishing, Boston (1993)
14. Hofstede, G.: Cultures and Organizations. McGraw-Hill Book Company Europe, Berkshire (1991)
15. Nocera, J., Hall, P.: Making sense in intercultural software implementation. In: Gunter, K., et al. (eds.) Culture and HCI: Bridging the Cultural and Digital Divides. Conference (6th: 2003). Proceedings of the 2nd British Computer Society Human Computer Interaction Workshop, University of Greenwich, London (2003)
16. Sheikh, J.A., Fields, B., Duncker, E.: Cultural Representation for Multi-culture Interaction Design. In: Aykin, N. (ed.) Proceedings of the 3rd International Conference on Internationalization, Design and Global Development: Held As Part of HCI international 2009, San Diego, CA, pp. 99–107. Springer, Heidelberg (2009), http://dx.doi.org/10.1007/978-3-642-02767-3_11
17. Fernandes, T.: Global interface design: A guide to designing international user interfaces. AP Professional, Boston (1995)
18. Sheridan, E.E.: Cross-cultural web site design: Considerations for developing and strategies for validating locale appropriate on-line content, MultiLingual Computing, Inc. (2001), http://www.multilingual.com (accessed 23/12/2003)
19. Hall, H.T.: Beyond Culture. Doubleday, New York (1976)

20. De Long, D., Fahey, L.: Diagnosing Cultural Barriers to Knowledge Management. Academy of Management Executive 14(4), 113–127 (2000)
21. Jarvenpaa, S.L., Staples, D.S.: Exploring Perceptions of Organizational Ownership of Information and Expertise. Journal of Management Information Systems 18(1), 151–183 (2001)
22. Gold, A.H., Malhotra, A., Segars: A Knowledge Management: An Organizational Capabilities Perspective. Journal of Management Systems 18(1), 185–214 (2001)
23. Baltahazard, P.A., Cooke, R.A : Organizational culture and knowledge management success: Assessing the behavior-performance continuum (Working Paper). Arizona State University West (2003)
24. Nonaka, Takeuchi, H.: The Knowledge-Creating Company. Oxford University Press, New York (1995)
25. Duncker, E.: Cross-cultural usability of the library metaphor. In: Proceedings of the 2nd ACM/IEEE-CS Joint Conference on Digital Libraries, International Conference on Digital Libraries Archive (2002)
26. Khoja, S., Sheikh, J.A.: KIIT Digital Library: An open hypermedia Application. In: International Conference on Use of Information Technology in Teaching of Science, Islamabad, Pakistan, March 12-14 (2002)
27. Romaine, S.: Language in Society. Oxford University Press, Oxford (1994)
28. Barrett, A.R., Edwards, J.S.: Knowledge Elicitation and Knowledge Representation in a Large Domain with Multiple Experts. Expert Systems with Applications 8, 169–176 (1996)
29. Zakaria, J.S., Stanton, J.M., Sarkar-Barney, S.T.M.: Designing and implementing culturally sensitive IT applications. Journal of Information Technology and People 16, 49–76 (2003)

Supporting of Requirements Elicitation for Ensuring Services of Information Systems Used for Education

Yuki Terawaki

The University of Tokyo Graduate School of Arts and Sciences, 3-8-1 Komaba, Meguro-ku, Tokyo 153-8902, Japan
Research Center for Computing and Multimedia Studies, Hosei University, 3-7-2 Kajinocho, Koganei, Tokyo 184-8584, Japan
terawaki@graco.c.u-tokyo.ac.jp

Abstract. This paper describes a method to ensure the quality requirements from service receiver in the requirement definition phase of system development. The proposed method measures the quality characteristics that are in the requirement document using the text-mining technique and concept dictionary and identifies requirements of document with quality characteristics of the International Standards Organization (ISO) / the International Electrical technical Commission (IEC) 9126-1:2001[1]. The case study shows that the quality characteristics are contained in the requirements document.

Keywords: Requirements Engineering; RE; Non-Functional Requirements; ISO/IEC 9126; text-mining approach.

1 Introduction

This paper describes a method to identify the service receiver's expectation of quality for developing system. Service receiver's needs are usually focused on not only the functional requirements to be met but also on quality characteristics, such as how soon the system can be run, how easy it is to use, or how often a system failure can occur etc. Such characteristics, collectively known as software quality characteristic or quality factors, are part of the system's non-functional requirements. It is widely recognized that in real systems, meeting the functional requirements often is more important than meeting the functional requirements in the determination of a system's perceived success or failure [11]. Therefore, in order to ensure services of developing system, the essential elements in the requirement definition are followings; The IT system-specific issues from various needs that were acquired from the stakeholders need to be selected. And then quality characteristics as much as possible need to be acquired and defined.

However, in the requirements acquisition phase of requirement definition, functional requirements are highly focused, non-functional requirements are not necessarily sufficiently defined. Moreover, some of the quality characteristics that are supposed to be fulfilled can occasionally be left unstated and implicit requirements because service receiver's requirements can be contradicted in requirements definition. Thus, to check

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 58–65, 2011.

the quality characteristics carefully is necessary for preventing of the implicit requirements but it also avoids a situation where users complain that the requirements are not all met when the developed system is delivered.

Therefore, this paper proposes a method measures the quality characteristics that are in the requirement document using the text-mining technique and concept dictionary and identifies requirements of document with quality characteristics of the ISO/IEC 9126-1.

The paper takes the following format. In section 2 describes the process of the requirements definition, and the non-functional requirements. In section 3 outlines the ISO/IEC 9126. The proposed method and its implementation are described in section 4. In section 5 introduces case studies. In section 6 conclusion and future works are provided.

2 Requirements Process and Non-functional Requirements

The RE Process has 4 processes. There is requirements elicitation, evaluation, specification (documentation) and quality assurance. This process is iteration on successive increments according to a spiral model [5,6]. A new iteration may take place at different stages of the software lifecycle. This spiral process model is fairly general and flexible. By participating both of provider and receiver (stakeholder) in the spiral process, requirements document will become elaborate. A stakeholder is a group or individual affected by the system-to-be, who may influence the way this system is shaped and has some responsibility in its acceptance. Stakeholders play an important role in the RE process. Because the problem of system embedded in human activity arises within some broader context. By stakeholder, requirements that has different viewpoint are elicited. Some of elicited requirements may include goal or problem on providing service. These are requirements against property of activity outside software. This is a non-functional requirement in the broad sense. Non-functional requirements define constraints on the way the software-to-be should satisfy its functional requirements or on the way it should be developed. Characteristics that fall into this category include how quickly it runs, how often it fails, and how easy it is to use.

Such characteristics, known as software quality attributes, as defined by ISO/IEC9126. Excellent software products reflect an optimum balance of competing quality characteristics. Quality attributes are difficult to define [10], because customers generally don't present their quality expectations explicitly. However in order to ensure services of developing information system, non-functional requirements should be well defined. Therefore, it is necessary to ascertain whether quality attributes were elicited during requirements definition.

3 ISO/IEC9126

ISO and IEC have developed the ISO/IEC 9126 Standards for Software Engineering – Product Quality [1-4] to provide a comprehensive specification and evaluation model for the quality of software products. The ISO/IEC 9126 standard is divided into four parts, under the general title Software engineering- Product quality:

- Part 1. Quality model
- Part 2. External metrics
- Part 3. Internal metrics
- Part 4. Quality in use metrics

The first three parts are concerned with describing and measuring the quality of the software product; the fourth part evaluates the product from the user point of view. We focus on part 1 (ISO/IEC9126 -1) which defined the quality model for external and internal quality.

ISO/IEC 9126-1 describes two-part model for software product quality: a) internal quality and external quality, and b) quality in use. The first part identifies the quality of a software product through six quality characteristics, namely: Functionality, Reliability, Usability, Efficiency, Maintainability and Portability. Each characteristic is subdivided into related sub-characteristics. These sub-characteristics are manifested externally when the software is used as a part of a computer system, and are a result of internal software attributes.

ISO/IEC9126-1 can be used not only the purpose to evaluate the quality of the developed software but also to define the non-functional requirements. Moreover, ISO/IEC 9126-1 also is used for infostructures' design. Examples of uses of the quality model are followings:

- Validate the completeness of a requirements definition;
- Identify software design objectives;
- Identify software testing objectives;
- Identify quality assurance criteria;
- Identify acceptance criteria for a completed software product.

4 Proposed Method

This method analyzes the rate of content of the quality characteristics in the requirements documents such as Request For Proposal (RFP). This method is able to calculate rate of content of whole document and each statement. The overview is described in Fig 1.

4.1 Morphological Analyzer and Index Generator

This method consists of two tools: one is the morphological analyzer (MA) and the other is the index generator (IG). MA breaks down the requirements specifications into sentences, and then each sentence is separated by morphological analysis. After that, MA tally the number of word and frequency of appearance, then makes term-document matrix.

In IG has some files. Each file include statements that represented the quality characteristics of ISO / IEC 9126. ISO / IEC 9126 consists of six quality characteristics so that IG also has six files. Each file of IG will contain the statements reviewed by specialists. IG first performs a morphological analysis against each file and removes characteristic words. Then, IG creates the frequency file. When frequency file are created, synonyms are added by using the concept dictionary. The reasons for adding synonyms are as follows. For example the word "user" is able to represent "customer" "End-user", or "consumer". Moreover The Japanese word for "user" is "riyousya" but "u-zer" is also recognized as Japanese. These meanings are identical, and such synonyms need to be considered within proposed method. After that, IG creates Word Article Matrix (WAM) file by executing mkw commands of Generic Engine for Transposable Association (GETA) [8]. Finally, this method analyzes the rate of content of the quality characteristics in the requirements documents using the term-document matrix in MA and the WAM file created in IG.

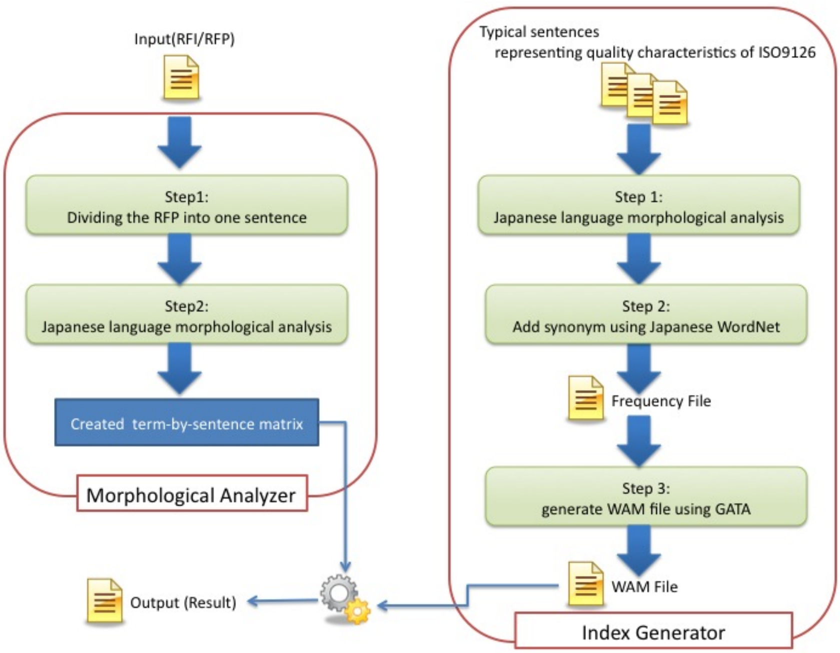


Fig. 1. Conceptual Diagram

4.2 Implementation

This method was developed as a CUI application by using Java programming language and shell scripts. The Japanese morphological analyzer Sen [7] is employed for morphological analysis in MA and IG. Synonyms are added to IG through Japanese WordNet [9], a concept dictionary, and extracted through a method that first acquires a lower level of the word group against the original words and then acquires the sum of the sets of each upper level word group. For the WAM file creation in IG and search operations in MA, a generic engine for transposable association (GETA) is employed. In IG, the mkw GETA command is used in the script. Since the search parts in GETA are provided only as the C library, GETA must create an execute format to wrap the I/O for the connection by using the Java program and standard I/O. The results will be output in CSV format, which facilitates easy use of the scores included in the search results.

5 Case Study

5.1 Overview

This section describes the trial practice using the proposed method. The RFP used for this trial practice was created in the university. This RFP involves replacing the network infrastructure and getting software (web-based application) across the university. In 2006, this RFP was already created by their system implementation committee after multiple reviews. However, author of this RFP is not the expert in Requirements Engineering (RE). In short, author didn't know the ISO/IEC 9126.

In this trial, the section relating to software, particularly Learning Management System (LMS) was picked out the RFP. This portion has about 4,600 characters.

Initially, This portion was analyzed manually by two experts of RE. After that it was analyzed by proposed method.

5.2 Result

Fig2 shows result of analysis by manually. Fig3 shows result of analysis by proposed method. The tendency of whole document denotes the same tendency. In this RFP, usability and functionality are describing most frequently requirements. However, portability was less described.

On the other hand, Fig4 shows quality characteristics contained in each sentence. For example, sentence of No.27 consists of Functionality, Reliability, and Usability. In their quality characteristics, Functionality is most remarkable tendency. Sentence of No.52 consists of Functionality, Usability. Functionality has much more higher than Usability.

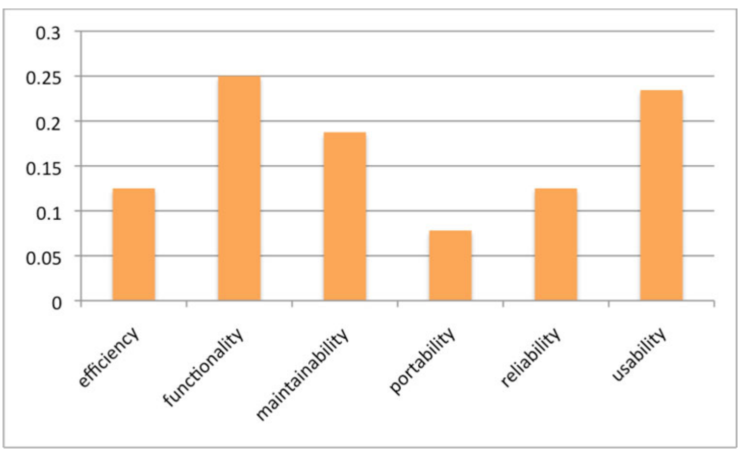


Fig. 2. Result of analysis by manually

5.3 Discussion

The results of analysis by manually and by proposed method were compared, and then these tendencies were similar. Let's take a look at each sentence Fig 4.

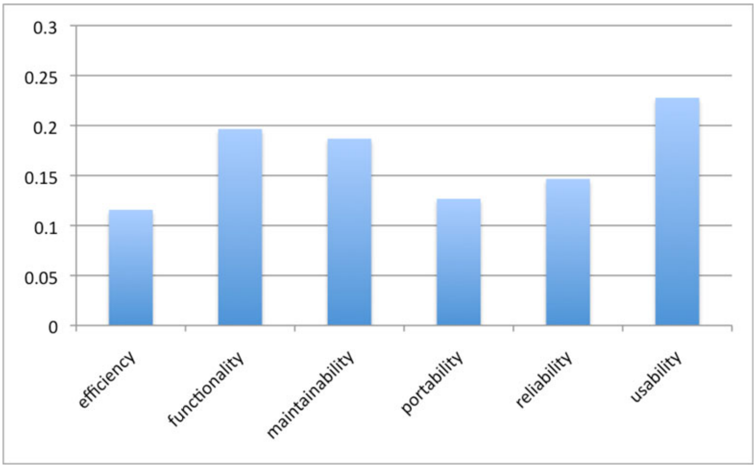


Fig. 3. Result of Analysis by Proposed Method

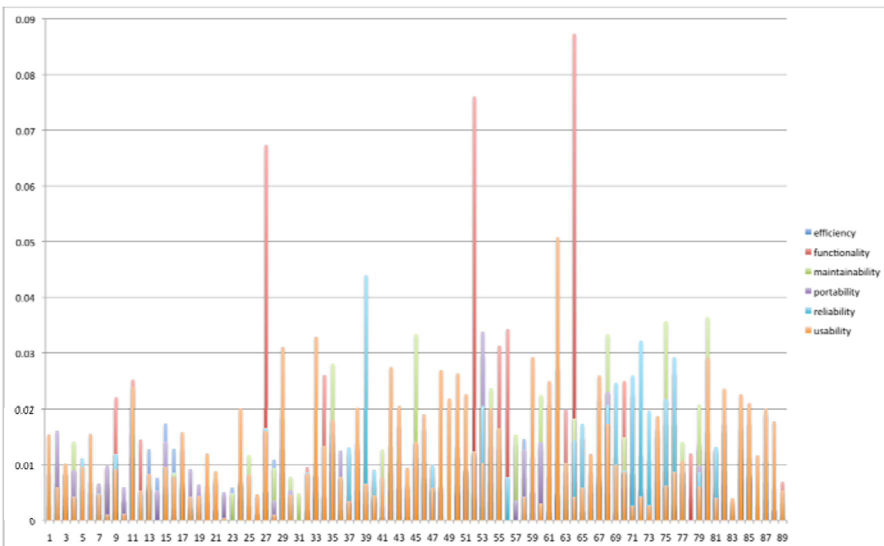


Fig. 4. Quality Characteristics in Each Sentence

For example, sentence of No.27 and No. 52 has features of Functionality. The requirement of No.27 is “The teaching material corresponding to international standard SCORM1.2 or more should be able to be published.” Manually powered analyzing determined that the requirements were represented Functionality. The requirement of No. 52 is ” When there is a difference in the method for SSO between LMS and University, method of SSO for LMS should be customized.” This requirements was deemed Functionality. This result also matched the analysis by manually. There are great examples. When two results are different, statement of target and the frequency file in IG need to be closely examined as to how different features were produced. At this time, when two results are different, we found that it is often the case that requirements sentence is subjective. These sentences supplies during requirements elicitation some clues about what they have in mind. On the other hands, some previous studies pointed out quality characteristics of ISO/IEC 9126 has ambiguity and omission [12]. Thus, Which characteristics of ISO9126 are most crucial to project success, we need to carefully consider.

6 Conclusion

In this paper, we illustrated a method and its implementation to ensure quality requirements from service receiver in the requirement definition phase of system development. The quality characteristics that are in the requirement document can be measured by using proposal method. Result of case study, the proposed method shows mostly good performance. Several issues are found through the case study in this paper.

- Remove some specific words

In any specifications there are some specific words that users wish to avoid considering as a morpheme. (E.g. 「情報メディア教育研究センター」 is 「情報:メディア:教育:研究:センター」). As such, it is not appropriate that there are some words, which are supposed to be treated as specific words, that are broken down into morphemes and reflected in the output of this framework. Therefore, we would like to review some words further for removal as specific words.

- Acquire synonyms

Though we acquire synonyms from the Japanese WordNet, acquiring them with no restrictions may create gaps between the original meanings. Currently, the method acquires lower level word groups against the original words and then acquires the upper level word group for each word in the lower level groups to determine the sum of the sets in the upper level word group. The top ten synonyms are then added. The number of synonyms added should be determined by continuously conducting actual case studies. Alternatively, the desirable number should be determined by examining the previous studies in the text-mining field.

Quality attributes are difficult to define, yet often they distinguish a product that merely does what it's supposed to from one that delights its user. However, service receivers generally don't present their quality expectations explicitly. Quality, in its many dimensions, must be defined from some clues about what they have in mind. In order to ensure service, the visible characteristics may be important. By gaining a lot of case study, the proposed method will be refined and then we'd like to learn which quality characteristics are most important to ensure service of information systems used for education.

References

1. ISO/IEC, ISO/IEC 9126-1 Software engineering- Product quality- Part 1: Quality model (2001)
2. ISO/IEC, ISO/IEC 9126-2 Software engineering- Product quality- Part 2: External metrics (2002)
3. ISO/IEC, ISO/IEC 9126-3 Software engineering- Product quality- Part 3: Internal metrics (2002)
4. ISO/IEC, ISO/IEC 9126-4 Software engineering- Product quality- Part 4: In Use metrics (2002)
5. Boehm, B.W.: A Spiral Model of Software Development and Enhancement. IEEE Computer 21(5), 61–72 (1988)
6. Kotonya, G., Sommerville, I.: Requirements Engineering: Processes and Techniques. John Wiley & Sons Ltd., Chichester (1997)
7. Sen (December 2010), `http://ultimania.org/sen/`
8. GETA (December 2010), `http://geta.ex.nii.ac.jp/e/index.html`
9. WordNet (December 2010), `http://nlpwww.nict.go.jp/wn-ja/index.en.html`
10. Bratthall, L., Wohlin, C.: Understanding Some Software Quality Aspects from Architecture and Design Models. In: 8th IEEE International Workshop on Program Comprehension (2000)
11. Charette, R.N.: Applications Strategies for Risk Analysis. Hill Software Engineering Series (1991)
12. Al-Kilidar, H., Cox, K., Kitchenham, B.: The Use and Usefulness of the ISO/IEO 9126 Quality Standard. In: International Symposium on Empirical Software Engineering (2005)

Visualizing Programs on Different Levels of Abstractions

Jo-Han Wu and Jan Stelovsky

1680 East-West Road,
Honolulu, HI 96822 USA
{johnwu,janst}@hawaii.edu

Abstract. To facilitate comprehension in an educational environment, a program should be ideally visualized using numerous depictions that employ different perspectives and levels of detail. Object-oriented programs that contain multiple levels of abstractions as a result of the modular design and encapsulation are primary candidates for such visualizations. ProViz is a framework that supports efficient creation of high-level visualizations for programs. Therefore implementing several visualizations of a program with ProViz has become a manageable effort. This article describes how a stack program is visualized on different levels of abstractions as an example of how multiple visualizations can be used in a course to help students understand data structures and algorithms.

Keywords: abstraction, object-orientation, software visualization, program visualization, algorithm.

1 Introduction

Visualization has been an invaluable tool for achieving program comprehension. In particular, high-level, storytelling type visualizations that use abstract representations and animations to display the semantic content of program data and algorithms, can be very effective in achieving a high degree of program comprehension [1], especially in an educational setting. Ideally an instructor should show not only one but various visualizations of an algorithm and the underlying data structures as different abstractions and perspectives are bound to reinforce the students' comprehension.

An object-oriented program contains layers of abstractions where each level conveys additional details of its inner workings[1]. Information hiding also leads to multiple abstractions as program details are encapsulated behind the public interfaces of classes. Correspondingly, different visualizations can reflect these levels of abstractions. For example, the abstraction of a calculator program from high to low can be: (1) behaving as a commonly-perceived calculator as the highest abstraction; (2) performing the computations using the underlying data structure, stack; (3) operating as an array or a linked list that constitutes a stack.

High-level visualizations are more suitable for initially understanding a task and for comprehension of high-level aspects of a program because they convey less detail

[1] While this also applies to other programming paradigms, this article solely focuses on the object-oriented paradigm.

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 66–75, 2011.

and let the learners focus on the main concept of the program. Low-level visualizations are more accurate, containing more details about the underlying implementation and thus are better for assisting in-depth analysis or debugging.

In programming practice, programmers need to develop a high-level mental model of how to solve a given problem. This mental model must be gradually broken down into a low-level conceptual model, which can then be translated into a programming language [4]. Teaching students to program is to follow the same scenario and guide them on the path of transforming mental models into low-level models within the programming domain. Visualizations matching the increasing complexity of the mental models can provide such guidance and help students transition from high-level concept to low-level mechanics with programming constructs. As a consequence, the students can first be shown a very abstract presentation that closely mimics the real problem domain and then be confronted with gradually more detailed visualizations that are closer and closer to the actual program code.

Until now, high, algorithm level of visualizations has been accomplished primarily using the *event-driven* visualization technique [2]. However, visualizing software programs on a highly abstract level is a time-consuming task [5]. Creating one, not to mention multiple, visual representation for a program can be laborious, and the costs of the visualization can easily outweigh its benefits.

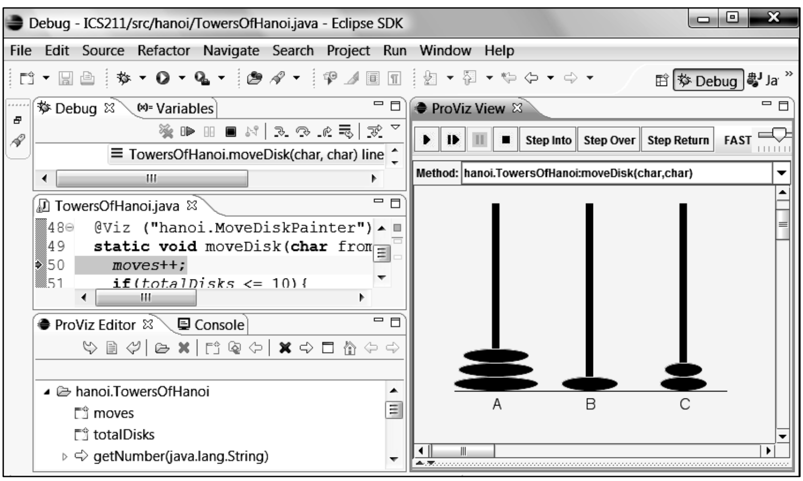


Fig. 1. A snapshot of ProViz in the Eclipse IDE

To facilitate the process of developing visualizations, we have developed ProViz, a program visualization framework that is embedded in the integrated development environment, Eclipse (as shown in Fig. 1). The ProViz framework allows an animator[2] to efficiently create high-level visualizations for programs. Using a customizable visual mapping mechanism, ProViz allows one target program to be associated with

[2] The role of an animator [1] is also known as a visualizer. Since the term is also used for visualizing data, "animator" will denote the author of the visualization in this article.

several visualization settings where each setting can depict various amounts of information and abstractions of the same program. In an educational setting, instructors can use ProViz to generate several visualizations of an algorithm or a data structure so that students can learn about it through multiple perspectives. Based on our experience in an introductory course, creating animation for programs which demonstrate the typical algorithms and data structures, such as searching, sorting, recursion, array operations, takes typically less than thirty minutes. Such comparably manageable effort has allowed us to create visualization for every sample program covered in the lectures or assigned as homework.

2 Visualizing Programs on Different Abstraction Levels

This section presents three different abstractions of a program that implements stack, the archetypical last-in-first-out data structure, and also shows how ProViz is used to visualize these abstractions.

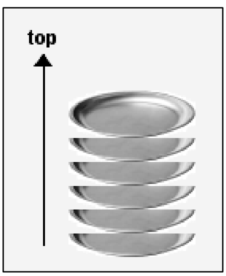


Fig. 2. A stack of plates

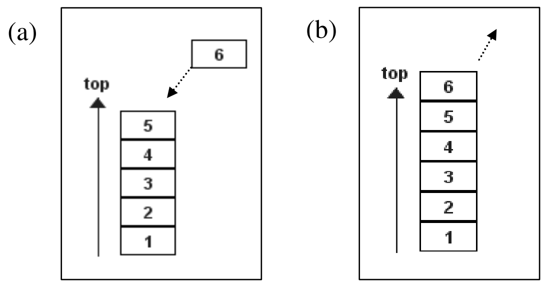


Fig. 3. Typical representation of a stack and: (a) the push method; (b) the pop method

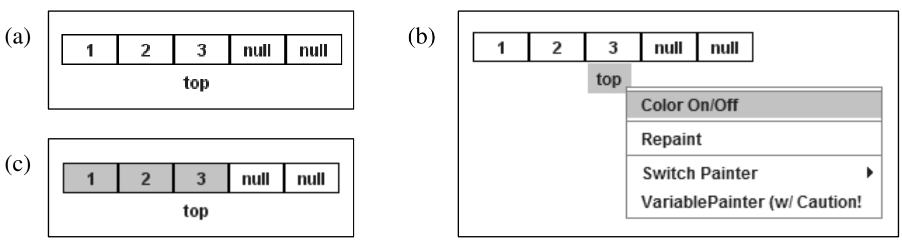


Fig. 4. Visualizing the underlying array of the stack with three numbers pushed. (a) In this example, the other elements in the array are null values. (b) Right-clicking on the "top" reveals custom actions defined by the painter for this index. (c) When the color option is switched on, the portion of the array occupied by the stack's elements is highlighted.

2.1 The Three Faces of Stack

Stack is often introduced with the example of a stack of plates in a dining hall. As such a presentation corresponds closely to the students' real life experience, it will be the first one that is introduce in the lecture (as shown in Fig. 2). The second variant shown in Fig. 3 depicts the expanding and shrinking data structure that mimics the behavior of the plates but shows the actual data. Since a stack can be implemented using an array as the underlying data structure, the last variant will show this array including its irrelevant elements (as demonstrated in Fig. 4a).

2.2 The Stack Program

One common way to implement a stack is to use an array as the underlying data structure and to keep an index that tracks the item at the top of the stack. The relevant skeleton of such a program in Java may be written as follows:

```
1   class Stack {
2     private String[] array = new String[5];
3     private int top = -1;
4     public void push (@DViz String element) {
5       if (top == array.length - 1) {
6         String[] newArray = new String [...];
7         ... //copies array to newArray
8         array = newArray;
9        } ... top++; ... }
10    public String pop () {...}
11    public static void main (String [] args) {
12      Stack stack = new Stack();
13      for (int i = 1; i <= 10; i++) {stack.push (i + "");}
14      for (int i = 1; i <= 10; i++) {
15        System.out.println(stack.pop()); }}}
```

The above program is defined by the `Stack` class. The `main` method starting on line 11 runs the actual program. It defines a `Stack` object and calls the `push` and `pop` methods several times to test the program.

To visualize the program amounts to visualizing the `Stack` object referenced by the variable `stack`. Thus the `stack` variable will be annotated with a *painter* – the ProViz's term for a Java class that draws a variable on the screen. At runtime, `stack` itself is in fact irrelevant to the behavior of the stack. It is the object's fields that truly define the stack, in particular the `array` field on line 2. The painter for `stack` will be a library painter in ProViz that does the underlying work by telling its field painter – the painter for `stack.array` – to perform the visualization. Therefore to create the three levels of abstractions which characterize the desired visualizations amounts to making different painters for the `array` field. This is elaborated in the following sections.

2.3 Visualization I - Typical Stack Representation

The first visualization demonstrated here is how the data structure "stack" is perceived by general programmers. As shown in Fig. 3, the stack merely consists of the elements that have been pushed into the stack, and the stack is presented as a vertical column anchored on its bottom edge and growing and shrinking at the top.

This abstract representation is achieved through the painter painting the `stack.array` field, which we will call `StackArrayPainter`. An array object actually consists of field elements with the names: `[0]`, `[1]`, `[2]`, and so on, and each of these fields will be automatically assigned a *field painter* of the `StackArrayPainter`. Therefore the first task of `StackArrayPainter` is to add to the screen its field painters from index `0` to the value of `top`. Secondly, this painter calculates and arranges the locations of its field painters so that they form the column with the painter for the element `[0]` on the bottom. As ProViz employs the event-listener model to let painters redraw their variable when its value changes, `StackArrayPainter` can listen to the `top` variable, and whenever `top`'s value changes, it adds or removes the necessary field painters. In addition, the visualization shown in Fig. 3 uses a painter for the `top` variable to draw the label "top" and the vertical arrow.

As a final touch, the `push` and `pop` methods can be visualized as well. When the `push` method is called, it animates moving the new element onto the top of the stack as shown in Fig. 2a. Similarly, the `pop` method moves the top element off the stack as shown in Fig. 2b.

2.4 Visualization II - Real-World Stack Representation

While the above visualization corresponds to how a stack is typically perceived by programmers, it may not emphasize the generic characteristic of a stack, in particular the understanding of how the push and pop operations affect one end of the data structure. Such a comprehension can be better achieved with a very high-level abstraction – the real-world example of stacking up plates. Such a representation allows the student to cognitively relate the real-world plates to the structure and operations of a stack – using the images of real plates.

This new visualization can reuse all the painters from Visualization I. The only change that is needed is to replace the field painters of `StackArrayPainter` by new painters that paint images of plates. This can be accomplished by a new `PlateImagePainter` class that inherits from ProViz's `ImagePainter` class. Implementation of this painter requires merely one line of code inside a single method that should return the path to the file with the plate image.

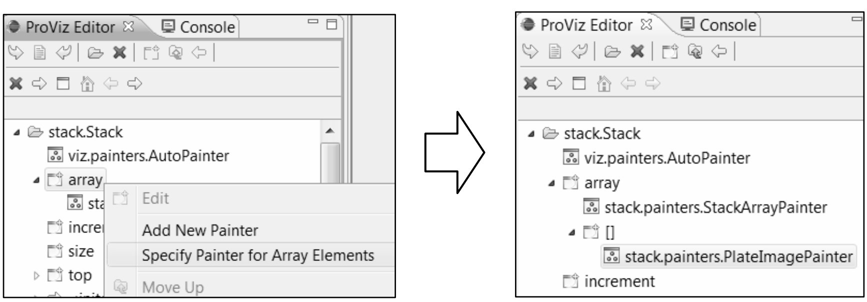


Fig. 5. Customizing the array's elements with an image painter by adding a painter for the field "`[]`" underneath the array

Once `PlateImagePainter` is developed, it needs to be associated with the fields of the `array` variable. This can be accomplished using *ProViz Editor*. Fig. 5 shows how a painter can be added to the array's fields (represented by the symbol "`[]`").

2.5 Visualization III - Uncovering the Implementation

Both of the above visualizations portray a stack corresponding to higher-level abstractions which hide the actual implementation of the program. To reveal how the stack is constructed, the last, low-level abstraction will visualize the program's data structures exactly as they are defined within the code.

First of all, the `array` variable will be visualized as an array in the default manner (accomplished using the @DViz annotation) Secondly, the `top` variable will be visualized by a new painter, `TopIndexPainter`, which inherits from `IndexPainter` in the ProViz library. Since `TopIndexPainter` acquired all the functionality of its parent class, it will position itself automatically under an array based on the value of `top`. `TopIndexPainter`'s only chore is to implement the `getArrayPainter()` method that returns the painter for the `array` variable.

As shown in Fig. 4b and 4c, `TopIndexPainter` can be more sophisticated and provide the option of highlighting the stack portion of the array, i.e. the elements up to the `top` index. Such coloration can effectively support the viewer's recognition of what part of the array currently represents the stack and which part is irrelevant.

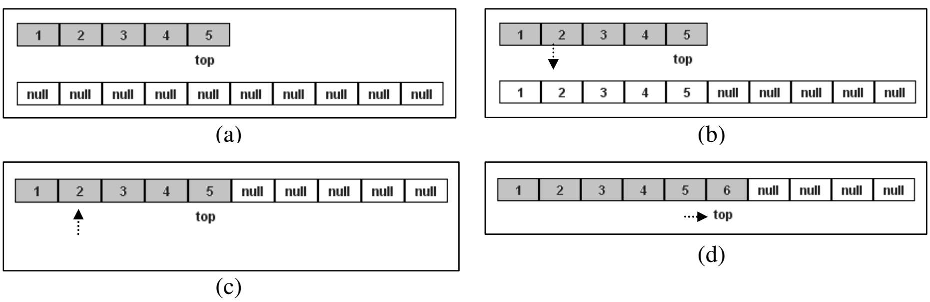


Fig. 6. Expanding the array: (a) a new array is allocated below the original array; (b) animation (indicated by the arrow) shows how the array's content is being copied; (c) the new array is moved up to the original location; (d) the new element is pushed and `top` is updated.

Array Expansion. When stack is implemented using a fixed-size array, the program needs to deal with a specific problem: when an item is pushed onto a stack and the array is full, the push method needs to allocate a new array and copy the contents of the old array into it (lines 5 to line 8 in the program code). This can be visualized by attaching a painter to `newArray` on line 6. In addition to drawing the new array, this painter can compose animations for the copy and replace actions as shown in Fig. 6.

3 The ProViz Framework

The ProViz framework is implemented as a plug-in within the Eclipse integrated development environment. It is written in Java and visualizes Java programs at the time of writing.

3.1 Visual Mapping

ProViz adopts the *data-driven* approach [2] to generating visualizations. This approach is also known as *state mapping*, since the program's data structures are mapped to visualizers. And as the program executes, changes in the states of the data cause updates of the visualization. The mapping between data and visualizers is the *visual mapping*. In ProViz, the *painters* are the visualizers, and therefore the visual mapping maps program elements, such as variables and methods, to painters.

In order to create high-level visualizations, the visual mapping in ProViz needs to be constructed by an animator. In contrast to automatic visualizations, which can only visualize program either in low abstraction or in fixed visual representations, visualizations at high levels of abstraction cannot be created automatically, as automation is known as the inverse of abstraction [1]. Therefore human authoring is required to develop customized painters and to construct the visual mapping. The main goal of a program visualization framework is thus to reduce the effort required in the process.

3.2 Viz Annotations

To define the visual mapping in ProViz, the animator uses *Java annotations* to specify (1) what program elements she wants to visualize, and (2) which painters will be used to visualize these elements. ProViz defines two annotations – @Viz and @DViz. The animator can use ProViz Editor to manipulate the visual mapping and also to save it to XML files. If the animator uses only this editor and the external XML files, she can eliminate Viz annotations altogether and avoid any code intrusion.

The @Viz annotation takes a list of painter names: *@Viz ("p1, p2, p3, …")*, e.g.:

```
@Viz ("DemoPainter, StringPainter[3]") Demo demo;
```

The first painter in this list, *p1*, will visualize the annotated element initially. If the user desires another depiction of an element, she can simply change the order of painters in ProViz, either within the annotation or at runtime by dragging a painter to the top with the mouse in ProViz Editor as shown in Fig 7.

The @DViz annotation declares that the variable should be visualized according to its type, provided that such class is annotated with a painter. If this class is not annotated, ProViz will its default painter to visualize this variable.

3.3 Painters

A painter is a Java class that inherits from the predefined `Painter` class. Creating a new painter therefore amounts to writing a subclass of an existing painter. ProViz maintains a continuously expanding library of painters built to visualize the semantics

[3] The painter class names shown here are shortened – the actual names must be fully-qualified.

of specific program elements. Thanks to this object-oriented design, new painters not only inherit necessary functionality from the top `Painter` but also require less code to customize. As the painter library expands, this reusability and customizability will reduce the work required for making new painters, thus making the effort of producing visualizations in ProViz more and more efficient.

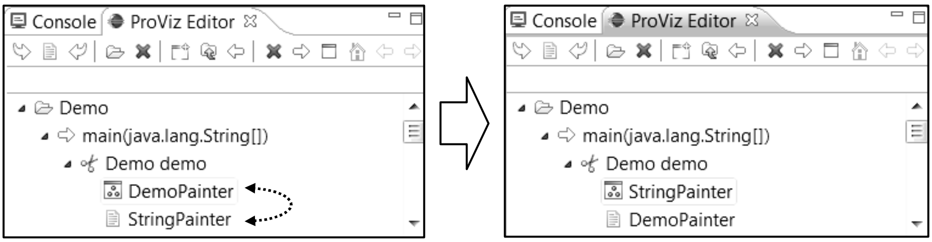


Fig. 7. Changing the order of painters in ProViz Editor

4 Discussion

As the level of detail increases the abstraction level decreases. Visualization II, on the highest-level of abstraction, is the easiest to understand as it uses the real-life model, focuses on the LIFO concept, and avoids all details, even the values of the elements. The more concrete Visualization I shows just about what programmers need to know when they use a stack. Lastly, the low-level Visualization III depicts the fact that the stack was implemented using an array and an index variable. The coloration shows the viewer the portion of the array that the stack occupies, reducing the cognitive effort of interpreting the top index as the boundary of the actual stack.

In the classroom, the high-level Visualization II was used to first introduce the concept of stack. Then Visualization I demonstrated that the stack's elements can contain values and still behave like a stack of plates. Visualization III showed the students how an array can be used to implement a stack and how an index tracks the top of the stack in the array. At last, the visualization of the array expansion can reveal the limitation of fixed-sized arrays and how the program overcomes it. Furthermore, the visualizations can be recorded and played back by ProViz's player. The students can not only review these recordings but also explore them interactively using the same controls the original visualizations have.

While Visualization I needs four painters, including method painters for push and pop, the Visualization II takes only one additional image painter as painters from Visualization I are reused. Visualization III uses two painters (one for a top index and a customized painter for the new array). The largest class has approximately 100 lines of code; all other painter classes are less than 60 lines of code, including many lines just for closing braces.) Each of the painters can be composed in the matter of minutes.

5 Related Work

Because creating high-level visualizations for programs is no easy task, dynamic program visualization is often limited to the display of low-level data or the display of the dynamic program data [8], such as call graphs, object and sequence diagrams, etc. High-level visualizations depicting the behaviors and algorithms of programs, aka *algorithm visualization*, is often handcrafted and not generated from a software program. ProViz bridges the two areas as a program visualization tool that is capable of generating algorithm visualizations.

Tarraingím [7] is a visualization environment for the SELF object-oriented programming language that was used for visualizations of data structures in different abstractions in [6] and [7]. Tarraingím produces visualization by monitoring public method calls rather than by observing the actual data. This unique approach may, however, exhibit problems with encapsulation and aliasing, as private methods cannot be visualized, and resolution of aliasing requires the compilation of owner trees [6].

jGRASP [3] is a lightweight programming environment for Java whose Object Viewer can visualize objects and data structures in common high-level depictions. As its visualizations are produced automatically, it does not visualize the high-level behavior of programs and does not support any explicit customizations for the user.

The Leonardo programming environment [2] is a data-driven visualization framework. To accomplish the visual mapping, calls to visualization routines are inserted as comments into the source code thus causing considerable code intrusion. While these calls do not alter the behavior of the program, the visual mapping is rigid and not easily customizable.

The most obvious way of producing program visualization is to identify interesting events in the program, compose graphical routines that visualize these events, and insert them into the source code at corresponding points. This "*event-driven*" approach was first adopted in BALSA and then followed by several other visualization systems [2]. While this concept is simple, the problem lies in its application. Identifying interesting events in the program is not easy and this approach is therefore mostly confined to small programs. Secondly, it modifies the program and causes code intrusion [2]. Therefore despite of its popularity, the event-driven approach may not be an ideal candidate for visualizing general programs.

6 Conclusion

The data-driven ProViz framework for constructing visualizations makes it possible to develop multiple perspectives of a program with comparably reasonable effort. The approach of visualizing programs in different abstractions has been very effective in teaching students in a CS2 course, as students were presented with step-by-step, gradually more detailed levels of visual abstractions to guide them towards more profound understanding of algorithms and data structures, starting from the high-level concept and down to the low-level implementation in the source code.

ProViz is capable of visualizing arbitrary Java programs as a result of a flexible visual mapping mechanism for composing different depictions as well as the object-oriented approach to the development of painters which encourages code reuse. It therefore can significantly reduce the effort required for making new visualizations.

Our future research will include (1) investigation of how visualizations can be optimized to support better transition from the conceptual model to the actual program code; (2) development of comprehensive and more sophisticated painters for common data structures; (3) visualizations of a comprehensive set of algorithms on multiple levels of abstractions. Moreover, a visual editor for the composition of new painters can further simplify the task of creating visualizations.

References

1. Cox, K.C., Roman, G.-C.: Abstraction in Algorithm Animation. Technical report WUCS-92-14, Department of Computer Science, Washington University (1992)
2. Demetrescu, C., Finocchi, I., Stasko, J.T.: Specifying algorithm visualizations: Interesting events or state mapping? In: Diehl, S. (ed.) Dagstuhl Seminar 2001. LNCS, vol. 2269, pp. 16–30. Springer, Heidelberg (2002)
3. Hendrix, T.D., Cross II, J.H., Barowski, L.A.: An Extensible Framework for Providing Dynamic Data Structure Visualizations in a Lightweight IDE. In: 35th SIGCSE Technical Symposium on Computer Science Education, pp. 387–391. ACM, New York (2004)
4. Jerding, D.F., Stasko, J.T.: Using Visualization to Foster Object-Oriented Program Understanding. Technical report GIT-GVU-94-33, Georgia Institute of Technology (1994)
5. Mukherjea, S., Stasko, J.T.: Applying Algorithm Animation Techniques for Program Tracing, Debugging, and Understanding. In: Proceedings of the 15th International Conference on Software Engineering, pp. 456–465. IEEE Computer Society, Los Alamitos (1993)
6. Aumann, Y.: Visualising Objects: Abstraction, Encapsulation, Aliasing, and Ownership. In: Diehl, S. (ed.) Dagstuhl Seminar 2001. LNCS, vol. 2269, pp. 58–72. Springer, Heidelberg (2002)
7. Noble, J., Groves, L., Biddle, R.: Object Oriented Program Visualisation in Tarraingim. Australian Computer Journal 27(4), 138–149 (1995)
8. Price, B., Baecker, R., Small, I.: An Introduction to Software Visualization. In: Stasko, J., Domingue, J., Brown, M.H., Price, B.A. (eds.) Software Visualization: Programming as a Multimedia Experience, pp. 3–27. The MIT Press, Massachusetts (1998)

Measurement and Evaluation in Service Engineering

Sakae Yamamoto[1], Miki Hiroyuki[2], and Hirohiko Mori[3]

[1] Tokyo University of Science, Faculty of Engineering,
11-3 Kagurazaka, Shinjuku-ku, Tokyo, 162-8601, Japan
[2] Oki Electric Ind. Co., Ltd.
[3] Tokyo City Univ.
sakae@ms.kagu.tus.ac.jp, hmiki@cf.netyou.jp, hmori@tcu.ac.jp

Abstract. This paper discusses how to advance the service engineering research. The service engineering is still established as study, and has not been completed. However, it dared to take up how to advance the service engineering research. As for the reason, the following three points are thought. First of all, the methodology of an original service engineering cannot establish it yet. The secondly, researchers involved in the service engineering have taken an active part to the research in various fields. Thirdly, the service engineering is researched therefore based on each one's current knowledge. It has been thought that it is necessary to devise a method that manages to be unified. Of course, what respect you should note while thinking about the system of the service engineering now is described. Especially, it is thought that the finding of a deeply related field to the Ergonomics is needed in the service engineering.

Keywords: Service Engineering, Kano Theory, Industrial Engineering.

1 Introduction

1.1 Service

Up to now, it had been used with service in the field that was called the third industry used by economics. Moreover, it has been used in the field like the service-producing industry and the service industry, etc. because of the industry specification classification. Service relates to IT, the software industry, and the net industry recently. When service is defined there, it is "The one that hardware and software can be offered to the user is served". It will be accepted by a lot of people that it is this definition. However, I want to discuss the relation of the essence of the service and the viewpoint of Ergonomics and HCI.

1.2 What Is the Service Seen from the Ergonomics?

First of all, the usability and accessibility are thought as for the service seen from the ergonomics. It is the same as providing of the product and the system with service for the user to achieve the usability and accessibility. Therefore, this service means it is guaranteed that the product is easy-to-use.

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 76–81, 2011.

The concept of service is taken up in the standard of ISO/IEC Guide 71: 2001 and ISO 9241-20:2008 in ISO. These standards are intended to help developers enable ICT equipment and services (and forthcoming novel or innovative equipment and services) so that they can be used by the widest range of people, regardless of their capabilities or disabilities, limitations or culture. In addition, these standards are declared for service to be provided the user in using hardware and software.

The thing offered to the user by the viewpoint of the ergonomics indicates usability and accessibility. Furthermore, Universal Design that is these broader concepts is indicated. In short, it thinks whether make it the product of the offered thing every the user.

An important thing is to make the engineer understand the idea of providing service for the product related to IT. The viewpoint of a current engineer of making the preferable product that the user wanted for the user was insufficient. Especially, they were not making the product by the viewpoint "It is possible to use it" for the person where the user had the senior citizen and the trouble. Then, the standard such as Guide71 and ISO9241-20 has come out.

It is thought the following two points are large and it is related to the background in which it has come to relate the usability and accessibility to service.

1. First of all, the advancement of ICT (Information Communication Technology) is given. That is, ICT cannot be separated from the advancement of the computer. The computer advances and has infiltrated our society. This brought various changes to our society. For instance, the word ubiquitous society was made to appear. However, what happiness did this society bring to our life? The advantage has not been shown yet as a concrete image. This made us expect new service to appear. In addition, the big role that this service plays has come into view. This gives the society a big impact, and the expectation that creates a further product and the system is becoming a reality. This also has not only manufacturing but also the influence on the information-technology industry. In addition, the expectation to which the user also receives a big favor is rising.
2. To see the favor to the user concretely, the design that reflects the user's favor and opinion and demand is requested. The user's reproof or demand relates to this design. Moreover, there is a technical improvement, and, in addition, user requirements evolve, too. In a word, contents of service always are evolving, too.

It thinks about consideration to the elderly person and the widest range of people who have not received the favor up to now based on these movements.

2 Service Engineering

The idea of services of the present age is different from the conventional mentality. The source of this modern idea is an idea called SSME (Service Sciences, Management and Engineering) that United States IBM advocated in 2004.

The service engineering is a research field in which it thinks how to achieve the service that has been described in the above-mentioned. It was described that the idea like the usability and accessibility was related to service in the viewpoint of the Ergonomics in the above-mentioned. It is to present the method of materializing the usability and accessibility to the product and the system. Not only the ordinary user but also the elder person and the disorder person can use the product and the system in recent years, and service that aims mastering it is thought ideally. What information is useful for that? The creation of service that is new and seems wonderful is requested.

2.1 Notes of Practice of Service Engineering

There is an overarching point though it has not been pointed out so much up to now by the service engineering research. It is that service changes. It is necessary to create service while understanding this respect. This idea is the same as the idea of the attractive quality method developed by Dr. Kano.

We were watching television, and for instance, it went to the place of the T.V. set of the television and we had changed the channel when we wanted to change the channel before. And, this is thought that we are natural, it was accustomed to this action, and this action has become natural.

However, it will end even if the channel tuning in and the loudness control by remote control can be done when a remote controller is developed, and it doesn't go to the T.V. set of the television separately. This becomes accustomed to this very conveniently. Then, remote controller becomes a natural function.

It connects it in remote control in addition, and it is possible to think, and to connect it to DVD and the network by putting various functions the hope of the user. This becomes an attractive function for the user.

The change in function of such a product very often occurs in our surroundings. For instance, a mobile phone is the same. The functional alteration of the product will change the content of the service provided for us like this. Therefore, it is important that the designer understand the service change. In addition, the change is closely related to user requirements.

Of course, it is a problem to accept user requirements without regard to principles. Moreover, it is not easy to reflect user requirements from a technical viewpoint. A newer requirement appears when the requirement is met. As for this, it is necessary to think about the change in service. This suggests that there be various levels in service.

This means the Dr. Kano's theory used by the quality control applies also in the service engineering. That is, the finding of the quality control is not to be able to disregard it in the service engineering. If service is considered to be a quality of course, there is not a sense of incompatibility and either it is likely to be accepted.

It is not said that any service engineering and the quality engineering are the same here.

Certainly, the improvement of the quality gives the user large benefit. In addition, the favor will change the usage of products in daily life and working. That is, not only the voice call function of the cellular phone but also E-mail (sending and receiving of the character) function and photograph function were added. It not only used an individual function but also a new usage became possible. For instance, shopping seeing the photograph of the commodity from carrying became possible. This shows that the form of shopping changed. Up to now, it has been shown that the user who has not shopped by telephone can shop even by carrying. Of course, it is shown that the sales form greatly changes into shop managers. In a word, the change in the quality function influences user and managers' purchasing. That is, the method of providing service changed. This is not only a story of a quality improvement alone.

In addition, it is appearance of the result of expanding, and using the influence to its maximum. The difference between the service engineering and the quality control of aiming was shown.

3 Object

The possible application of the theory in fields other than the service engineering is discussed, and the methodology for that is brought together.

3.1 Level of Service (Three Levels of Kano Theory)

According to Dr. Kano, the quality is constructed by three stages, and he explained these as three levels. The first level is called the "Must-Be quality", and the evaluating point is that the products are satisfying the basic function or not satisfying. It is guaranteed to talk over the telephone anywhere if it is a cellular phone.

The second level means the customers' satisfaction, and this level was considered after the competitor was able also to have the "Must-Be" function.

The quality at the first level is achieved, and the customer requirement develops the product to fill it afterwards. An important evaluation point is how it was able to meet the demand at that time. Therefore, the second level is defined as a satisfaction rating of the specified demand of the customer.

The third level indicates the state that cannot compete only by customer satisfaction with the other companies. The designers need to recognize the customer's potential

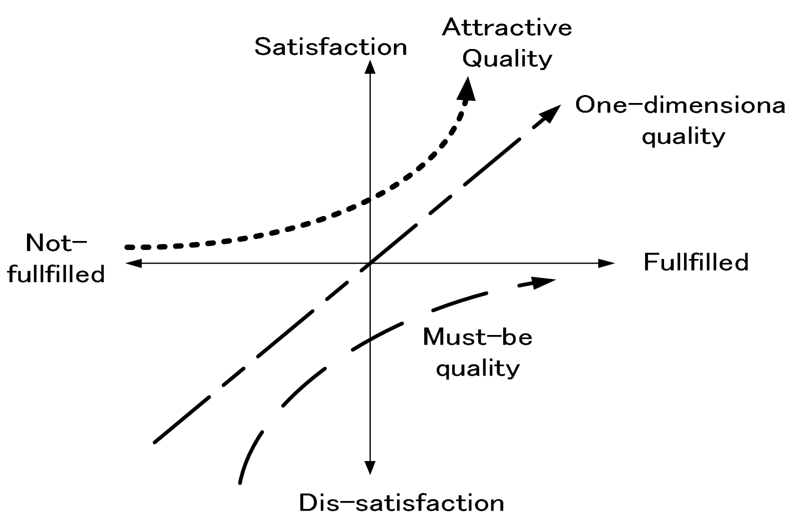


Fig. 1. The three levels of Kano theory (Figure of Kano (1) was modified)

demand. If designer make the product which satisfied the customer's potential demand, the customer accept the product with pleasure that exceeds the customers' expectation. It can be said that the achievement of customer delight. These three levels quoted Dr. Kano's classification. Kano calls the first level Must-be Quality, calls the second level, and One-Dimensional Quality and the third level are called Attractive Quality.

4 Meaning of Change in Service

The reason for the factor to cause the change in service is that user requirements change as described in above-mentioned in "Level of service". Catching the change in this demand mentioned that the change in service would be caught above. Then, in order to examine the change in user requirements, the questionnaire will be useful means. Whenever each level is achieved, do we research using the questionnaire? Of course, though then, we need check between the demands and the technical development.

Then, when a unified quality is invented from the must-be quality without taking the questionnaire again when user requirements are filled, it is likely to have to think about the following of that. Thinking the structure between each factor is clarified by analyzing the factor to invent a unitary quality for that is important.

5 Evaluation of Service

It demands generally from the user in the questionnaire survey. At this time, it is necessary to clarify scale of assorted traits. That is, the statistical technique that can be used according to each standard at the same time as distinguishing nominal scale, ordinal scale, interval scale, and ratio scale is different. When it makes a mistake in this, a significant result is not obtained. When it makes a mistake in this, a significant result is not obtained. Moreover, you should check this because there is a method of taking out the key word of a free description, too recently. How to summarize data has not been solved yet. It is easy for the principal component analysis and the factorial analysis to interpret it when bringing it together from three at most in four.

6 Cognitive Method and Industrial Engineering (IE) Method

When the user's behavior is examined, the action analysis and the motion study are often used. Especially, there is a purpose of examining it from cognitive viewpoint. Therefore, a mere action is not analyzed but the reason for the action why man acted like this is clarified. As a result, the operation process when man uses the product can be clarified. Thus, it becomes possible to clarify the operation that cannot use the product well, and to examine the reason.

A necessary thing is how to find out the reason for the manipulatory behavior by questioning from the operator here. Then, verbal protocol method is used. As a result, it becomes possible to clarify the reason for the operation, and to clarify the background how user requirements have come out. The persona method and the scenario method are used recently. These are called cognitive viewpoint. However, a basic technique often uses the IE method. When the IE method applied to the field, there were many failed cases in the past. Actually, the means restrict the target, and it fails. Users are devising the method of IE because we are borrowing the IE technique as an analysis technique now that the purpose is different. However, human analyze by means of IE method. It worries about whether poke one's head inside to the above-mentioned failure unconsciously.

The current state of IE today has changed. The application only of the technique is not thought. The problem solving on the field is a purpose of IE. It thinks whether this problem solving becomes a new service creation by the service engineering. When the technique is used, a meta idea might be needed. Especially, what service develop by you while thinking about the change in service is demanded now.con

7 Peroration

It took a general view from a peripheral study area where the service engineering was surrounded of the service engineering. And, it thought about the influence given to the service engineering by discussing the possible application of peripheral study.

Acknowledgements. This paper was described around having noticed in the discussion with Dr. Noriaki Kano. Dr. Kano taught me the hint of the Kano theory and service engineering, and I touched the essence of Kano theory moreover. I wish to express my gratitude deeply for here.

If my paper made digression from the Kano theory, I had the responsibilities in this.

References

1. Kano, N., et al.: Attractive Quality and Must-Be Quality. The Journal of Japanese Society for Quality Control 14(2), 39–48 (1984)
2. Kano, N.: Quality crisis that Japan faces. DAIWA Review (9), 4–13 (2003)
3. Stauss, B., et al. (eds.): Services Science, Fundamentals, Challenges and Future Developments. Springer, Heidelberg (2008)
4. Salvendy, G., Karwowski, W.: Introduction To Service Engineering. John Wiley & Sons, New Jersey (2010)

A Human Interface Toolkit for Developing Operation Support System of Complex Industrial Systems with IVI-COM Technology

Yangping Zhou[1,*], Yujie Dong[1], Xiaojing Huang[1], and Hidekazu Yoshikawa[2]

[1] Institute of Nuclear and New Energy Technology, Tsinghua Univerisity, Energy Science Building, Haidian District, Beijing 100084, China
Tel.: +86-10-62783555, Fax: +86-10-62771150
[2] Kyoto University/Harbin Engineering University, Nantong Street, Nangang District, Harbin 150001, China
{zhouyp,dongyj,huangxj}@tsinghua.edu.cn,
yosikawa@kib.biglobe.ne.jp

Abstract. A human interface toolkit is proposed for helping the user to develop operation support system of complex industrial system such as Nuclear Power Plant (NPP). With a friendly graphical interface, this integrated tool includes a database, a procedure editor and a procedure executor. A three layer hierarchy is adopted to express the complexity of operation procedure, which includes mission, process and node. There are 10 kinds of node: entrance, exit, hint, manual input, detector, actuator, data treatment, branch, judgement and plug-in. The operation support system will sense and actuate the actual industrial systems with the interface based on IVI-COM (Interchangeable Virtual Instrumentation-Component Object Model) technology. A prototype system of this human interface toolkit has been developed is applied to develop a simple operation support system for a simulated NPP.

Keywords: Human Interface Toolkit, Operation Support System, Complex Industrial System, IVI-COM.

1 Introduction

Because of the development and maturation of computer and related technologies, the digitization is inevitably happening in many fields of complex industrial system such as Nuclear Power Plant (NPP) [1]. It is believed that the application of these digital operation support systems is able to improve the safety and reliability of complex industrial system and reduce the worker's work load [2-5]. However, the design, development and maintenance of operation support system such as digital operating procedure under both operational states and accident conditions require not only a profound understand of design, operation and structure of NPP but also expertise on information technology.

* Corresponding author.

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 82–89, 2011.

Because of the reasons mentioned above, a human interface toolkit is proposed for helping the user to develop the operation support system of complex industrial system. With a friendly graphical interface, this integrated tool includes a database, a procedure editor and a procedure executor.

In this database, a three layer hierarchy is adopted to express the complexity of operation procedure, which includes mission, process and node. There are 10 kinds of node: entrance, exit, hint, manual input, detector, actuator, data treatment, branch, judgement and plug-in. With the procedure editor, user can easily develop and maintain the procedure and the finished procedure will be stored in the database. Then, the procedure executor can load the procedure from the database for operation support and thus act as a digital operation support system. The operation support system will sense and actuate the actual industrial systems with the interface based on IVI-COM (Interchangeable Virtual Instrumentation-Component Object Model) technology [6] embedded in detector node and actuator node. With the help of various nodes, processes and missions, the developed digital system can access information from plant, make interaction with operator, call additional application, and so on.

According to the design mentioned above, a prototype system of this human interface toolkit has been developed with Visual C++, DirectX SDK, MSXML Parser SDK, Microsoft Agent SDK and SQL server. The prototype system is applied to develop a simple operation support system for a simulated NPP. Several operating conditions are tested on the prototype system and the developed operation support system in order to validate and improve the proposed work.

2 Hierarchy for Operation Support

A three layer hierarchy is adopted to express the complexity of operating procedure, which includes mission, process and node. A mission denotes a specific task which the operator wants to finish, such as drop of the power during normal operation, startup of plant, etc. Usually, a mission can be divided into several independent processes. A certain process consists of a chain of actions, here, it called as nodes. At present, ten kinds of node are applied to express the various actions in the operating process:

- Entrance: Startup point of whole mission or entrance from upstream process to present process.
- Exit: End point of whole mission or exit from present process to downstream process.
- Hint: Can provide information to operator with text or video.
- Manual input: Operator can manually input a value to the support system.
- Detector: Provide input interface with a real detector in an industrial system based on IVI–COM (Interchangeable Virtual Instrument) technology.
- Actuator: Provide input interface with a real actuator in an industrial system based on IVI-COM technology.
- Data treatment: Treat the value in support system with numeric calculation, string treatment, base conversion, etc.
- Branch: Automatically determine the destination node according to the judgement conditions.

- Judgement: Operator can make a logical (YES/NO) judement according to the condition.
- Plug-in: Provide interface with external program which can start the external program if necessary.

There are two kinds of connectors in this hierarchy, process connector and nod connector. The process connector connects the downstream process with the upstream process. The node connector connects the downstream node with the upstream node. In addition, the process connector is affiliated with an exit in an upstream process and an entrance in a downstream process at node level.

Fig. 1 shows a simple example of this hierarchy. There are four processes in mission 1: process 1, 2, 3, and 4. There are four nodes in process 1: entrance 1, hint 1, detector 1 and exit 1. And four nodes in process 2: entrance 2, hint 2, detector 2 and exit 2. The process connector between process 1 and process 2 is affiliated with exit 1 in process 1 and entrance 2 in process 2 at node level. Sometimes there is more than one process connector between two processes but only one node connector between two same nodes.

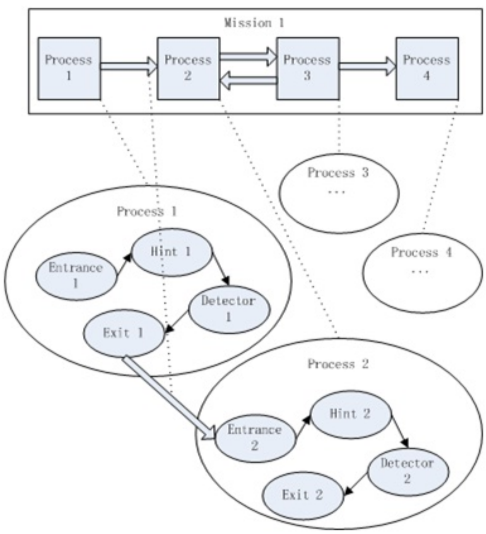


Fig. 1. Hierarchy for operation support

3 Design of Procedure Development Toolkit

The relations between target system, operating procedure and procedure development toolkit are shown as Fig. 2. The procedure development toolkit has a procedure editor and a procedure executor. The work of procedure development toolkit consists of two stages, editing stage and executing stage.

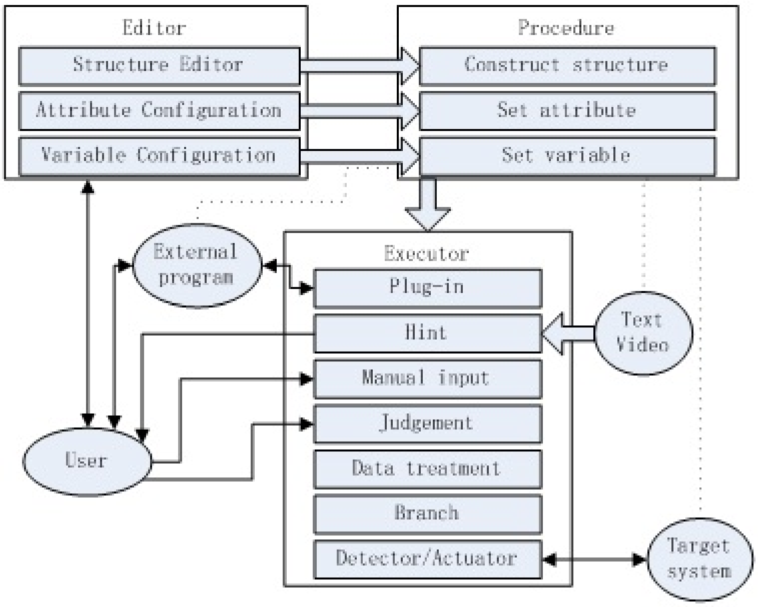


Fig. 2. Target system, operating procedure and procedure development toolkit

During editing stage, the operating procedure will be constructed and maintained by user with the help of the editor. The external multimedia file such as text and video will be affiliated with the hint node in the procedure to provide information to user. In order to extend the ability of developed support system, a plug-in function is provided in the plug-in node which is embodied with an interface for loading the external program and transferring data between the developed system and external program. The attribute of detector and actuator node can be set with the editor to configure IVI-COM interface between procedure system and actual detector and actuator. A set of variable can be established with the help of the editor, and the name, type and value can be set by the variable configuration tool. In addition, the variable can be affiliated with the various nodes. With a friendly graphical interface, user can establish the digital procedure whose information including mission, process, node, connector and variable is stored in a database system.

During executing stage, the executor loads the procedure from the database and acts as an operation support system. When a mission of the procedure is activated, the executor will start from an entrance node which is a startup point of the mission. Then, the working point of the procedure will go to next node which is the downstream node of this entrance node. The executor will act differently according to the type of the node. If the node is a plug-in, the affiliated external program will be activated, which can load needed data and provide support to operator. The hint node will provide information to operator by using affiliated text and video. Through manual input node, user can input the value to the variable of system. When arriving at judgement node, user can make a YES/NO answer to the question. Data treatment node can treat the value of variable and branch node can make a jump to a certain node according the value of a variable. Detector node will get value from the real detector in the target system and actuator node will actuate the real actuator in the target system for system control. The detector node and the actuator node work based on IVI technology.

4 IVI Architecture

IVI is an integral component of a National Instruments test system. The IVI Architecture is shown as Fig.3. IVI sits above the VISA I/O layer in the program hierarchy and is integrated into the application development environments provided by National Instruments. The IVI architecture breaks the traditional instrument driver into two parts – an instrument-specific driver and a class driver. The instrument-specific driver functions the way traditional instrument drivers have in the past, but with an underlying architecture that is optimized for performance and includes instrument simulation. The class instrument driver contains generic functions for controlling an instrument category and calls the corresponding instrument-specific driver functions at run time. The test program can be written with either the class driver or the specific driver.

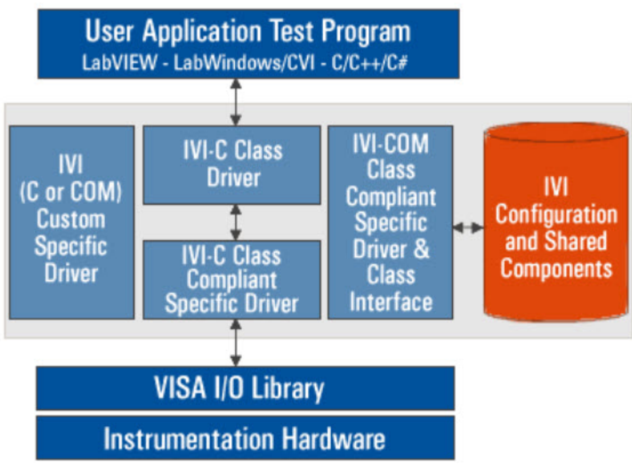


Fig. 3. IVI architecture

The IVI Foundation has defined the IVI architecture to work with two interface technologies, one based on the ANSI C standard (IVI-C) and one on Microsoft COM (Component Object Model) technology (IVI-COM). The two architecture types are designed to be interoperable.

5 Prototype System of Integrated Tool

According to the design mentioned above, a prototype system of this integrated tool has been developed with Visual C++, DirectX SDK, MSXML Parser SDK, Microsoft

Agent SDK and SQL server. The prototype system is applied to develop a simple operation support system for a simulated NPP. Several operating conditions are tested on the prototype system and the developed operation support system in order to validate and improve the proposed work.

Fig. 4 shows the snapshot when the procedure editor of prototype system is working to establish the demo operating procedure. The explorer can manage the mission, processes and nodes in a tree list mode. User can easily find, choose or delete the object in the procedure with various operations such as expand, fold, etc. With the help of mission editor, user can append, delete and choose the symbols of process and process connector with simple mouse operation. User can conveniently append, delete and choose the symbols of different nodes and node connectors with the process editor. At the bottom of the window, configuration assistant can assist user to configure the attribute of mission, process, node, process connector and node connector when the relevant object is chose in explorer, mission editor or process editor. In addition, the variable can be added, deleted and configured in configuration assistant.

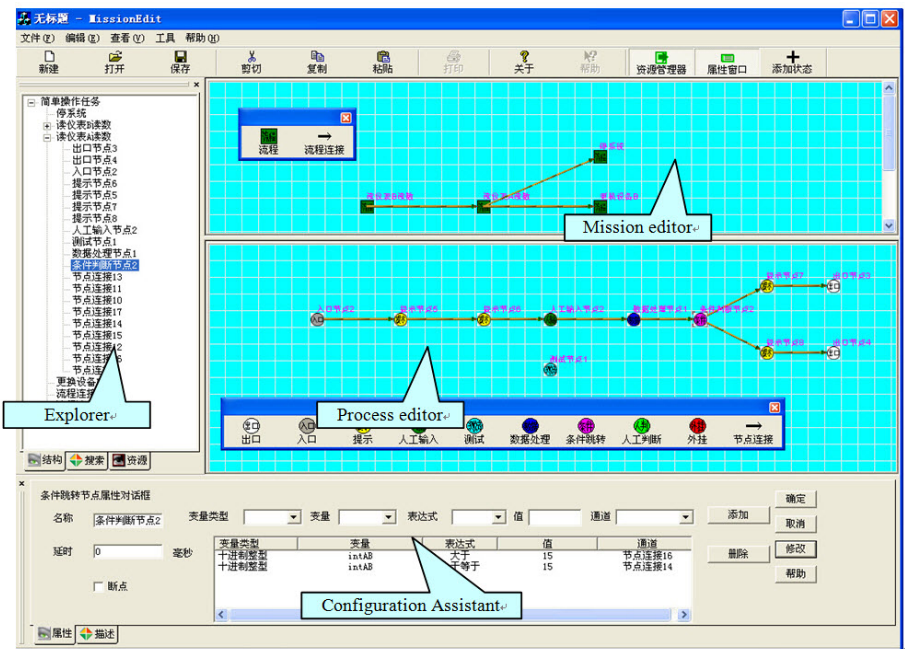


Fig. 4. Editor of procedure toolkit

Fig. 5 shows the scene when the procedure executor of prototype system loads the procedure and acts as an operation support system. The executor can work in three modes: debug mode, run mode with breakpoint and run mode without breakpoint. In the debug mode, the executor will stop at each object to help the developer of procedure confirm and validate the procedure. In the run mode with breakpoint, the executor will stop at breakpoint and interaction node. The executor will only stop at interaction node, such as manual input, judgement, etc, in which the interaction between user and procedure executor is needed. In the executor, the explorer can be used to manage the mission, processes and nodes in a tree list mode. The present node which the executor runs on will be marked in the explorer. In the interaction window, user can make interaction with the executor. For example, user can input a value to a variable when a manual input node is executed. The signal display can display the state of the signal which is derived from detector and related node. The variable display can be used to monitor the state of variable.

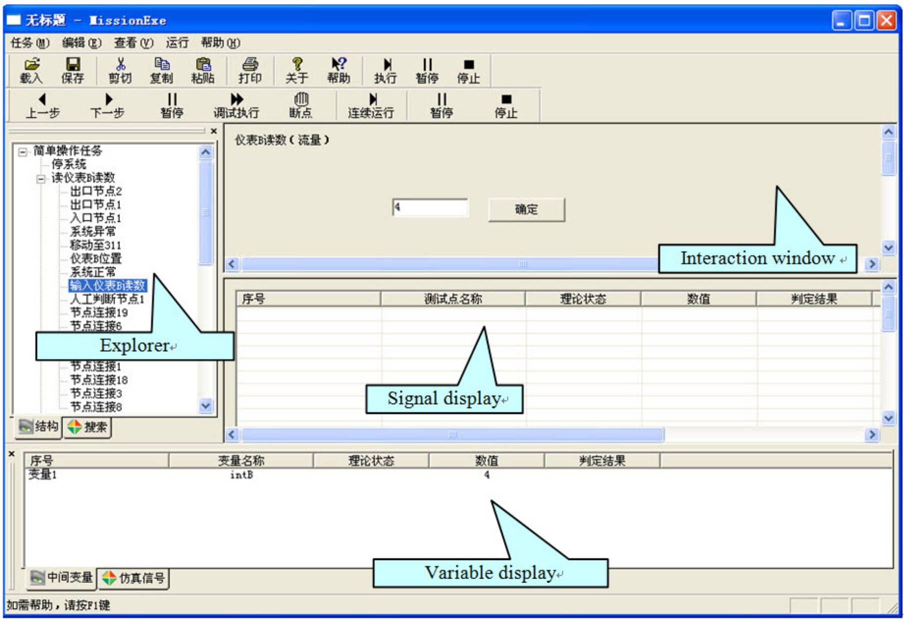


Fig. 5. Executor of procedure toolkit

6 Conclusions and Perspectives

With an integrated graphical interface, a procedure development toolkit, is proposed for the development and maintenance of application for Man-Machine Interaction (MMI) such as operation support system, etc. In terms of this integrated tool, operating procedure for different target systems can be constructed, maintained and implemented conveniently.

In this study, a prototype system of this integrated tool has been developed by using Visual C++, DirectX SDK, MSXML Parser SDK, Microsoft Agent SDK and SQL server. A demo operation support system for a Nuclear Power Plant simulated by RELAP5/MOD2 has been developed by using this prototype system. Several operating conditions are tested on the prototype system and the developed operator support system in order to validate and improve the proposed work.

The proposed integrated tool is now far away from the actual application to the industrial system. In future, the proposed integrated tool will be improved according to the following considerations:

- Possible support to smooth and rapid migration from paper based operating procedure to digital operating procedure.
- Improvement of the human interface especially for the executor by considering the actual requirement of operator when operating the industrial system.
- Interface for connecting other software and hardware in the actual control room efficiently and reliably.
- Improvement of reliability and efficiency of the integrated tool.

References

1. Zhang, L.: Design of Digital Operating Procedure for FuQing NPP. In: Technical Meeting on the Impact of Digital I&C Technologies on the Operation and Licensing of NPPs, Beijing, China (2008)
2. Liu, F., Zhang, Z., Peng, M.: Development of Emergency Operation Procedure System for Nuclear Power Plants. Chinese Nuclear Power Engineering 29(4), 104–108 (2008)
3. Park, J., Jeong, K., Jung, W.: Identifying cognitive complexity factors affecting the complexity of procedural steps in emergency operating procedures of a nuclear power plant. Reliability Engineering and System Safety 89, 121–136 (2005)
4. Niwa, Y., Hollnagel, E.: Integrated computerisation of operating procedures. Nuclear Engineering and Design 213, 289–301 (2002)
5. Lee, S.J., Seong, P.H.: Development of automated operating procedure system using fuzzy colored petri nets for nuclear power plants. Annals of Nuclear Energy 31, 849–869 (2004)
6. IVI information, http://www.ni.com/ivi/

Part II

Information and User Interfaces Design

A Conceptual Model of the Axiomatic Usability Evaluation Method

Yinni Guo[1], Robert W. Proctor[1,2], and Gavriel Salvendy[1,3]

[1] School of Industrial Engineering, Purdue University, West Lafayette, IN
[2] Department of Psychological Sciences, Purdue University, West Lafayette, IN, USA
[3] Department of Industrial Engineering, Tsinghua University, Beijing, China
{guo2,rproctor,salvendy}@purdue.edu

Abstract. This paper describes a new usability evaluation method developed for consumer electronics. This method, Axiomatic Evaluation, is based on Axiomatic Design theory, a formalized methodology that can be used to represent a variety of design problems. From this perspective, to generate a design that meets the perceived needs, designers should first specify the design goals of “what we want to achieve”, and then start the design process with a clear description of “how we will achieve it”, so that the recursive “design/build/test” cycle could be reduced.

Keywords: axiomatic design, usability engineering.

1 Introduction

Usability has been studied in Human-Computer Interaction (HCI) for about two decades to improve the efficiency and effectiveness of interactions. Several evaluation methods have been used to detect usability problems, mostly for websites and software but also for other products. However, some authors have argued that HCI usability methods are not well suited for application to consumer electronic products. Consequently, our objective was to develop and describe a new usability evaluation method for consumer electronic products.

2 Background Literature

2.1 Usability Evaluation Methods

Usability evaluation seeks the best way to improve the quality of systems or products in terms of effectiveness, efficiency, and satisfaction. Traditional usability evaluation methods, including cognitive walkthrough, heuristic evaluation, think aloud, interview, focus group, and questionnaire, are widely used to test computer-based products like software and web sites, and have met with good success. However, some authors suggest that the same concept of software usability cannot be employed effectively for consumer electronic products. For example, Han, Yun, Kim, and Kwahk stated, “It is not appropriate to apply the same concept of software usability developed in the

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 93–102, 2011.

HCI research directly to the consumer electronic products" [1]. One reason for this assessment is that there are several aspects of the interfaces and interactions with consumer electronic devices that are different from those of the computer-based products [2]. For computer-based products like web sites or software, users can use the keyboard or mouse to control the interface. These control tools, especially the mouse, can be moved freely and quickly all over the interface. Compared to these set- ups on the computer, control of the interfaces for consumer electronic products is limited and more difficult.

Compared to computer-based products like software and websites, consumer electronic products like mobile phones impose limitations such as smaller screen size and cumbersome input mechanisms [3]. Also, because the consumer electronic products have hardware-oriented features such as dedicated controls in addition to software-oriented features such as menus, there are more dimensions that need to be considered with regard to usability [4]. Unlike just clicking the mouse, users of consumer electronic products typically need to press the control panel or remote control several times when they want to change a setting or go to a certain function. Sometimes they cannot even determine how to operate the device because there is not enough information about control actions or the design of operation does not fit the users' mental models [1]. Users may encounter situations of pressing the button more times than needed and then have to spend additional time pressing the keys again. A typical example is how users change their television or monitor settings.

Because of these differences, researchers have tried to establish distinct usability guidelines for consumer electronic products. For mobile devices, guidelines include: five usability criteria for mobile UI design [5], a usability checklist based on 21 usability principles [6], and nine critical factors for handset's perceived usability [2]. Besides the studies of the devices, service for consumer electronic products has also been investigated to enhance the usability. It has been suggested that personalization is essential for creating a positive mobile experience [3, 7]. Similar to web-based products, the "fun" attribute contributes to users' adoption of wireless handheld devices, even more than the perceived usefulness does [8].

2.2 Axiomatic Design

Introduction. How the design process should be structured is always an issue. In the past, many engineers have designed their products interactively, empirically, and intuitively, based on creative insights from years of experience. Undisciplined designs have resulted that are not only poor in terms of customer satisfaction, but also with respect to missed schedules, cost overruns and failed products. Symptoms of poor design include low quality and unreliability, as well as high costs of hardware, software, and systems [9]. Some design errors are major problems that have been well publicized. However, there are also many "small" problems that simply inconvenience or aggravate consumers. All bad designs can be dangerous, cost money, limit the usefulness of products, or delay the introduction of new products.

To be efficient and to generate a design that meets the perceived needs, the designer must specifically state the design goals before the design process begins. Therefore, a rigorous design approach must start with an explicit statement of "what we want to achieve" and end with a clear description of "how we will achieve it" [10].

Once the customers' needs are understood, these needs must be transformed into a minimum set of specifications, or functional requirements, that adequately describe "what we want to achieve" to satisfy the customers' needs. The descriptor of "how to achieve it" may be in the form of design parameters.

Axiomatic design is a formalized methodology that can be used to represent a variety of design problems [9]. The ultimate goal of axiomatic design is to establish a scientific basis for design and to improve the expertise of the designer by providing a theoretical foundation based on logical, rational thought processes and tools. Axiomatic design theory is based on two fundamental axioms that minimize the possibility of making mistakes when products are being developed. The theory helps to overcome shortcomings of the product development process based on a recursive "design/build/test" cycle, which requires continuing modifications and changes as design flaws are discovered through the testing [11].

Axiomatic design consists of four domains and two axioms. The four domains, ordered in consideration, are: *customer* domain, *functional* domain, *physical* domain and *process* domain. An earlier domain in the series represents "what we want to achieve", whereas the subsequent adjacent domain represents the design solution of "how we propose to satisfy the requirements specified in the prior domain." The customer domain is characterized by customer attributes (CAs), or what the customer is looking for in a product, process, systems or materials. In the functional domain, the customer's needs are specified in terms of functional requirements (FRs) and constraints. To satisfy the specified FR, design parameters (DPs) in the physical domain can be conceived. Finally, to produce the product specified in terms of DPs, a process that is characterized by process variables (PVs) in the process domain can be developed. Axiomatic design activity consists of mapping from one domain to the other, and domains have different meanings in different research areas.

The mapping process between the domains can be expressed mathematically in terms of characteristic vectors that define the design goals and design solutions. At a given level of the design hierarchy, the set of functional requirements that define the specific design goals constitutes the FR vector in the functional domain. Similarly, the set of design parameters in the physical domain that have been chosen to satisfy the FRs constitutes the DP vectors. The mathematical expression can be written as [FR] = [A][DP], where [A] is called the design matrix that characterizes the product design. The design matrix is of the following form for a design that has three FRs and three DPs, This equation can be written in terms of its elements as equation 1 or 2:

$$\begin{aligned} FR_1 &= A_{11}DP_1 + A_{12}DP_2 + A_{13}DP_3 \\ FR_2 &= A_{21}DP_1 + A_{22}DP_2 + A_{23}DP_3 \\ FR_3 &= A_{31}DP_1 + A_{32}DP_2 + A_{33}DP_3 \end{aligned} \tag{1}$$

$$[FR] = [A]\,[DP] \tag{2}$$

$$[A] = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \tag{3}$$

where the two axioms in axiomatic design are identified by examining the common elements that are always present in good designs. The axioms are also identified by examining actions taken during the design stage that result in dramatic improvements.

The first axiom, called the *Independence Axiom*, states that the independence of FRs must always be maintained, where FRs are defined as the minimum set of independence requirements that characterizes the design goals. The Independence Axiom states that when there are two or more FRs, the design solution must be such that each one can be satisfied without affecting other FRs. That means that a correct set of DPs must be chosen to be able to satisfy FRs and maintain their independence.

The second axiom suggests that physical integration is desirable to reduce the information content if the functional independence can be maintained. This axiom is called the *Information Axiom*, and it states that among those designs that satisfy the Independence Axiom, the design that has the smallest information content is the best. To make the design work, the information amount must be estimated by the users. Since the information content is defined in terms of probability, the second axiom also states that the design that has the highest probability of success is the best design.

During the mapping process (e.g., from functional domain to physical domain), one must first make the correct design decisions using the Independence Axiom. When several designs that satisfy the Independence Axiom are available, the Information Axiom can be applied to select the best design. To satisfy the Independence Axiom, the design matrix must be either diagonal or triangular. When the design matrix is diagonal (e.g. [A] in equation 4), each of the FRs can be satisfied independently by means of one DP. Such a design is called an *uncoupled* design. When the matrix is triangular (e.g. [B] in equation 5), the independence of FRs can be guaranteed if and only if the DPs are determined in a proper sequence. Such a design is called a *decoupled* design. Any other form of the design matrix is called a full matrix and results in a *coupled* design, for which the end result will be conflicting control. Therefore, when several FRs must be satisfied, designs must be developed that will enable creation of a diagonal or triangular design matrix:

$$[A]=\begin{bmatrix} A_{11} & 0 & 0 \\ 0 & A_{22} & 0 \\ 0 & 0 & A_{33} \end{bmatrix} \tag{4}$$

$$[B]=\begin{bmatrix} B_{11} & 0 & 0 \\ B_{21} & B_{22} & 0 \\ B_{31} & B_{32} & B_{33} \end{bmatrix} \tag{5}$$

In a design problem, there can be many designs (different sets of FRs) that satisfy the Independence Axiom. However, one of those designs is likely to be superior. The Information Axiom provides a quantitative measure of the merits of a given design and is useful in selecting the best among the acceptable designs. In addition, the Information Axiom provides the theoretical basis for design optimization and robust design, stating that the design with minimum information content is the best.

According to this axiom, the design with the smallest I is best because it requires the least amount of information to achieve the design goal. Information content I_i for a given FR_i is defined in terms of the probability P_i of satisfying FR_i,

$$I_i = \log_2 \frac{1}{P_i} = -\log_2 P_i \tag{6}$$

The information content is expressed in bits of information. When all m FRs are statistically independent, the information content for the system is,

$$I_{sys} = \sum_{i=1}^{m} I_i = -\sum_{i=1}^{m} \log_2 P_i \tag{7}$$

When all m FRs are not statistically independent, the information content for the system (shown in equation 7) becomes,

$$I_{sys} = \sum_{i=1}^{m} I_i = -\sum_{i=1}^{m} \log_2 P_{i|\{j\}} \quad \text{for } \{j\}=\{1, ..., i-1\} \tag{8}$$

Applications. Axiomatic design has been applied to a variety of products and systems such as mechanical design, system design and control, software design, organizations management, materials design and more. Many innovations have been made and commercialized based on the use of axiomatic design, such as microcellular plastics and woven electrical connectors [11]. Although the meanings of the four domains are quite different from one product area to another, axiomatic design successfully enhances the design quality in different areas. As many case studies presented suggest [11], the performance, robustness, reliability and functionality of products, process, software, systems and organizations are significantly improved when the axioms are satisfied.

Applications to Usability Design. Ergonomic systems must be designed to be robust and efficient in satisfying their FRs and constraints [12]. Unless FRs related to ergonomics are defined at the beginning of the design process, it is difficult to incorporate human factors after the design is completed. It is equally difficult to modify an existing system or product to overcome ergonomic shortcomings. Therefore, it is important to “design it right from the beginning” [12]. Ergonomics has both time-independent and time-dependent aspects. Some products may not satisfy human factors from the very beginning of product introduction because of deficiencies in design, but even good products may deteriorate as they are being used [12]. To deal with the former, one must be certain that FRs related to human factors are incorporated and satisfied at the time of design.

In recent years, axiomatic design principles have been applied in human factors to ergonomic microscope workstation design [13], biomechanical hand tools [13], and in HCI to information visual design [14] and e-commerce websites [15]. However, there is not yet much research using axiomatic design to improve usability of human-computer/human-machine interfaces.

3 Conceptual Model of Axiomatic Evaluation Method

Traditional usability evaluation methods are more proper for use on software and websites than on consumer electronic products. Traditional usability evaluation methods may also neglect usability problems related to the complexity of a product, defined here as a measure of uncertainty in achieving the specified FRs. According to research [11], one type of complexity named imaginary complexity occurs because users lack the knowledge about the function or design. Complexity will cause more errors and require a longer time for users to operate a function. Therefore, by using the "axiomatic evaluation" method developed from the axiomatic design theory, it is possible to discover usability problems regarding complexity of design.

Axiomatic design is different from traditional design theories because it starts designing from "what we want to achieve" to "how we could achieve it", with the former listed as domain vectors and is the latter examined through the mapping matrix between two domains. The complete design process starts from "customer domain" to "functional domain", to "physical domain", to "process domain". Since the final goal of usability evaluation is to examine and enhance the design's efficiency, as well as effectiveness and customer satisfaction, it is meaningful and feasible to start the "mapping" process from "what do customers want" and proceed to "what does the product provide". The framework of axiomatic design is used to set up three domains: "Customer Domain", "Functional Domain" and "Control Domain" and the mappings between them. Table 1 compares the domains, mappings, axioms, and constraints for the axiomatic design and axiomatic evaluation method. The biggest difference in domain settings is that "Control Domain" is the 3rd domain instead of "Physical Domain". The main reason for this change is that as an evaluation process, the physical design variables or parameters cannot be obtained. What can be obtained easily are the control parameters, which are the control keys. The "Customer Domain" [CA] can be retrieved by an open-ended questionnaire. Participants can be provided a list of existing functions and asked what their expected functions are, or any extra expectations. The "Function Domain" [FR] and "Control Domain" [CT] can be gained by interviewing the designer as well as reading the product manual. The uses of the two axioms are similar, but the axiomatic evaluation method has extended constraints for information appliance design such as stimulus-response compatibility [16] and Hick's Law [17], according to which response time increases as a function of the amount of uncertainty among alternatives, or, the amount of information conveyed by an action.

Usability problems can be found by examining the mapping between the three domains. The mapping can be expressed as matrices [X] and [Y]:

$$[\text{Customer Requirement}] = [X]\,[\text{Function}] \tag{9}$$

$$[\text{Function}] = [Y]\,[\text{Control}] \tag{10}$$

If the [X] matrix has a "0" row, it means that the customer requirement could not be met, and the function needs to be added to that product if several participants have the same response. If the [X] matrix has a "0" column, it means that there are redundant functions, and we need to delete the function if all or most participants show the "0" column.

Table 1. Comparison of Axiomatic Design Theory and Axiomatic Evaluation Method

	Axiomatic Design Theory	Axiomatic Evaluation Method
1st domain	[CA]: Customer domain. The needs (or attributes) that the customer is looking for in a product or process or systems or materials.	[CA]: Customer domain. Customer requirements of a certain product.
2nd domain	[FR]: Functional domain. The minimum set of independent requirements that completely characterizes the functional needs of the product	[FR]: Functional domain. The existing functions of a certain product.
3rd domain	[DP]: Physical domain. The key physical variables in the physical domain that characterize the design that satisfies the specified FRs.	[CT]: Control domain. The control keys or elements of the product.
4th domain	[PV]: Processing domain. The key variables in the process domain that characterize the process that can generate the specified DPs.	Not defined.
Independent Axiom	The independence of functional requirement (FRs) must always be maintained. The design solution must be such that each one of the FRs can be satisfied without affecting other FRs. That means we have to choose a correct set of DPs to be able to satisfy the FRs and maintain their independence.	The independence of functions must be maintained. Each function must be able to be executed by control keys without conflicting. Moreover, control keys for the same function should be grouped together.
Information Axiom	Reduce the information content if the functional independence can be maintained.	When the Independent Axiom is satisfied, this axiom is used to reduce the information amount, and reduce choices in decision making for users so that they could decrease reaction time and reduce error rate.
Constraints	No constraints for the general Axiomatic Design method. There are different constraints applied in various areas	Decision Making Hick-Hyman Law Stimulus-Response Compatibility

The mapping between [CA] and [FR] also provides an index of function sufficiency which can be used at the beginning stage of product development. If the average index is high, the current product could be said to satisfy most customers. Otherwise, the product needs to be added more functions.

The mapping between the Functional Domain and the Control Domain shows how easy it is to control the appliance. According to the "Independence Axiom", an ideal matrix should be an uncoupled matrix, or at least a decoupled matrix. An uncoupled matrix is a diagonal matrix, which means the users only need one control step, or one button to operate the function. However, most consumer electronic products have more than 20 functions, and it is difficult to have more than 20 buttons on the panel. Thus, several keys usually need to be pressed to accomplish a certain task.

A decoupled matrix is another option. This means that customers may need to press several keys to achieve their goal, but the keys should not conflict with other functions. Moreover, if the Information Axiom is satisfied, usability constraints for the control keys should also be met. For instance, according to action selection studies, keys for the same task should be grouped together. Therefore, by examining [Y], one can find out whether the Independence Axiom and usability constraints are satisfied.

$$\begin{bmatrix} \text{Make phone calls} \\ \text{View recent calls} \\ \vdots \end{bmatrix} = \begin{bmatrix} Y & \cdots & 0 \\ \vdots & \ddots & \vdots \\ Y & \cdots & Y \end{bmatrix} \begin{bmatrix} \text{Key 1 (green call key)} \\ \text{Key 2 (red stop key)} \\ \vdots \\ \text{key n} \end{bmatrix} \tag{11}$$

Usability problems can be found by recording the participants' operating action. If a participant's requirement for the cell phone is to send messages, s/he will be asked to finish the task of sending a message while their operating action is recorded and arranged to the matrix [Y]. If the [Y] matrix is a decoupled matrix, then it means there are no conflicting controls between different functions. If it is not a decoupled matrix, the designers should redesign the controls. However, only meeting the aim of a decoupled matrix is not enough. The designer needs to make sure that users do not need to make too many key presses, and for the same function the keys should be grouped together.

The procedure of the axiomatic evaluation method should be used during the first stage of development. Potential advantages of using the new evaluation method are:

a) Only a few experts are needed. One or two experts knowledgeable about usability and axiomatic evaluation would be enough.
b) A large number of participants may not be needed. Choosing participants that represent the composition of the customers would be sufficient.
c) Extra problems can be discovered. The current usability evaluation method of Heuristic Evaluation is based on the "guidelines", and Cognitive Walkthrough is based on "a designed task". These usability evaluation methods might lead to usability interface problems but neglect other types of usability problems between the interface and other parts of the device such as the control panel. Axiomatic evaluation is more related to customers and would discover the problems that are of concern to them. Examining the mapping between [CA] and [FR] would show what the missing content is (that customers need) and what the redundant content

is (that is barely needed by customers). By examining the mapping between [FR] and [CT], one would know what the problems are between the content/function and operation of the content/function.

d) Direct solutions for the discovered usability problems related to controlling can be given. One solution is to rearrange the use of keys or buttons and make the [Y] as a less coupled design to fulfill the Independence Axiom. Another solution is to reduce the number of choices to fulfill the Information Axiom. Other possible solutions include rearranging the keys on the panel so that the stimulus-response compatibility principles apply.

4 Follow-Up Study and Conclusions

Based on this conceptual model, an experiment was designed to test the axiomatic evaluation method [18]. Sixty people were randomly assigned to a group using a think aloud method and a group using used the axiomatic evaluation method. Three popular consumer electronic devices representing different levels of complexity were tested by each participant. More usability problems with the cell phone were found by the axiomatic evaluation method than by the other methods. Across all three products, the axiomatic evaluation method was better in finding problems of user requirements. These results suggest that traditional usability methods are not sufficient for evaluating complex devices like smart phone and GPS. Instead, the axiomatic method should be used.

References

1. Han, S.H., Yun, M.H., Kim, K.-J., Kwahk, J.: Evaluation of Product Usability: Development and Validation of Usability Dimensions and Design Elements Based on Empirical Models. International Journal of Industrial Ergonomics 26, 477–488 (2000)
2. Zhang, T., Rau, P.-L.P., Salvendy, G.: Exploring Critical Usability Factors for Handsets. Behaviour & Information Technology 29(1), 45–55 (2010)
3. Venkatesh, V., Ramesh, V., Massey, A.P.: Understanding Usability in Mobile Commerce. Communications of the ACM 46(12), 53–56 (2003)
4. Kim, J., Han, S.H.: A Methodology for Developing a Usability Index of Consumer Electronic Products. International Journal of Industrial Ergonomics 38, 333–345 (2008)
5. Ciavarella, C., Paterno, F.: Design Criteria for Location-aware, Indoor, PDA Applications. In: Chittaro, L. (ed.) Mobile HCI 2003. LNCS, vol. 2795, pp. 131–144. Springer, Heidelberg (2003)
6. Ji, Y.G., Park, J.H., Lee, C., Yun, M.H.: A Usability Checklist for the Usability Evaluation of Mobile Phone User Interface. International Journal of Human-Computer Interaction 20(3), 207–231 (2006)
7. Ho, S.Y., Kwok, S.H.: The Attraction of Personalized Service for Users in Mobile Commerce: an Empirical Study. ACM SIGecom Exchanges 3, 10–18 (2002)
8. Bruner, G.C., Kumar, A.: Explaining Consumer Acceptance of Handheld Internet Devices. Journal of Business Research 58, 553–558 (2005)
9. Suh, N.P.: The principles of design. Oxford University Press, New York (1990)

10. Lee, D.G., Suh, N.: Axiomatic Design and Fabrication of Composite Structures: Applications in Robots, Machine Tools, and Automobiles. Oxford series on advanced manufacturing, p. 451. Oxford University Press, US (2006)
11. Suh, N.P.: Axiomatic design: advances and applications. Oxford University Press, New York (2001)
12. Suh, N.P.: Ergonomics, Axiomatic Design and Complexity Theory. Theoretical Issues in Ergonomics Science 8(2), 101–121 (2007)
13. Helander, M., Lin, L.: Axiomatic Approach in Ergonomics Design with an Extension of the Information Axiom. Journal of Engineering Design 13(4), 321–339 (2002)
14. Quill, L.L., Batchelor, C., Kancler, D.E., Revels, A.R.: Application of Information Visualization Principles at Various Stages of System Development. In: Proceedings of 45th Human Factors and Ergonomics Annual Meeting, pp. 1713–1717 (2001)
15. Yenisey, M.M.: Axiomatic Design Approach for E-commercial Web Sites. LNCS, pp. 308–315. Springer, Heidelberg (2007)
16. Proctor, R.W., Vu, K.-P.L.: Selection and Control of Action. In: Salvendy, G. (ed.) Handbook of Human Factors and Ergonomics, 3rd edn., pp. 89–110. John Wiley, Hoboken (2006)
17. Schneider, D.W., Anderson, J.R.: A Memory-based Model of Hick's Law. Cognitive Psychology 62, 193–222 (2011)
18. Guo, Y., Zhu, Y., Salvendy, G., Proctor, R.W.: Axiomatic Usability Evaluation Method for Complex Information Appliances (under review)

Study on Evaluation of Kawaii Colors Using Visual Analog Scale

Tsuyoshi Komatsu and Michiko Ohkura

Shibaura Institute of Technology, 3-7-5, Toyosu, Koto-ku,
135-8548 Tokyo, Japan
l07057@shibaura-it.ac.jp

Abstract. In the 21st century, the Kansei values of industrial products are considered very important. Kawaii is one of the important Kansei values for future interactive systems and industrial products. However, since only few studies have focused on kawaii attribute, we systematically analyze the kawaii interfaces themselves, that is, the kawaii feelings caused by such attributes as shapes, colors, and materials. In previous experiments, we obtained interesting tendencies about kawaii attributes. For example, if an object has more brightness and more saturation, more participants chose it as most kawaii for every hue. The most commonly chosen hue was purple for both men and women. However, since we have not studied intermediate hues based on the Muncell Color System, we systematically experimented on the color elements to clarify the tendencies of kawaii colors. The experimental results clarified the tendencies of kawaii colors for each color element, such as hue, saturation, and brightness.

Keywords: kawaii, color, hue, saturation, brightness, vas, Kansei, virtual object.

1 Introduction

In the advanced information society of the 21st century with its communication infra-structure of computers and networks, software must be enhanced that utilizes this digital content. Among such content, Japanese kawaii characters including Hello Kitty and Pokemon have become popular all over the world, suggesting that the power of Japanese cuteness can be attained worldwide [1]. Kawaii is one of the important Kansei values for future interactive systems and industrial products. However, since few studies have focused on the kawaii attributes of the interfaces of interactive systems or other artificial products, we systematically analyze the kawaii interfaces themselves: kawaii caused by such interface attributes as shapes, colors, textures, and materials. Our aim is to clarify features of kawaii interface from the research results. We previously performed experiments and obtained interesting knowledge about kawaii attributes [2,3,4,5,6]. Our experiment concentrated on such color elements as hue, saturation, and brightness. We employed five basic hues based on the Muncell Color System (MCS). For each hue, we selected three connected

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 103–108, 2011.

values of saturation and of brightness. Thus, the total number of kawaii color candidates was 45. As an object's brightness and saturation increased, more participants chose it as most kawaii for every hue. The most commonly chosen hue was purple for both men and women. However, since we have not studied intermediate hues based on MCS, we addressed this problem in a new experiment. This paper describes the experimental results.

2 Experimental Set-Up

We showed the virtual objects on a 46-inch 2D/3D compatible LCD monitor from Hyundai on which participants watched stereoscopically with polarized glasses. For the kawaii shapes of objects, a torus was employed based on the results of our previous studies [2]. To select the candidates of kawaii colors, we used MCS. Color has three elements: hue, saturation, and brightness. We employed five basic hues (R, Y, G, B, and P) and five intermediate hues (YR, GY, BG, PB, and RP) based on MCS. For each hue, the following four colors were selected:

- #1 is white, which has the highest brightness and the lowest saturation.
- #2 has higher brightness and lower saturation than the base color.
- #3 is a base color with high brightness and high saturation.
- #4 is pure color, which has lower brightness than the base color and the highest saturation.

These colors for each hue were selected (Fig. 1). Because #1 for each hue is the same color, the total number of kawaii color candidates was 31. The background color was gray.

3 Evaluation

We used the Visual Analog Scale (VAS) to evaluate the kawaii degrees. For the evaluation of pain severity and relief, the method commonly used is the VAS [7]. Subjects arbitrarily marked 200-mm segments. The left side line doesn't seem kawaii, but the right side line does. The length from the left side to the mark put on segments by subjects is converted into scores from 0 to 100.

4 Experimental Methods

First, the color blindness of the participants was tested with the Ishihara color test. Next, they were shown four colors of the same hue and simultaneously evaluated the kawaii degrees of the four colors with VAS. This evaluation was repeated for each hue.

The ten hues were shown randomly. An example of the presented set of objects is shown in Fig. 2.

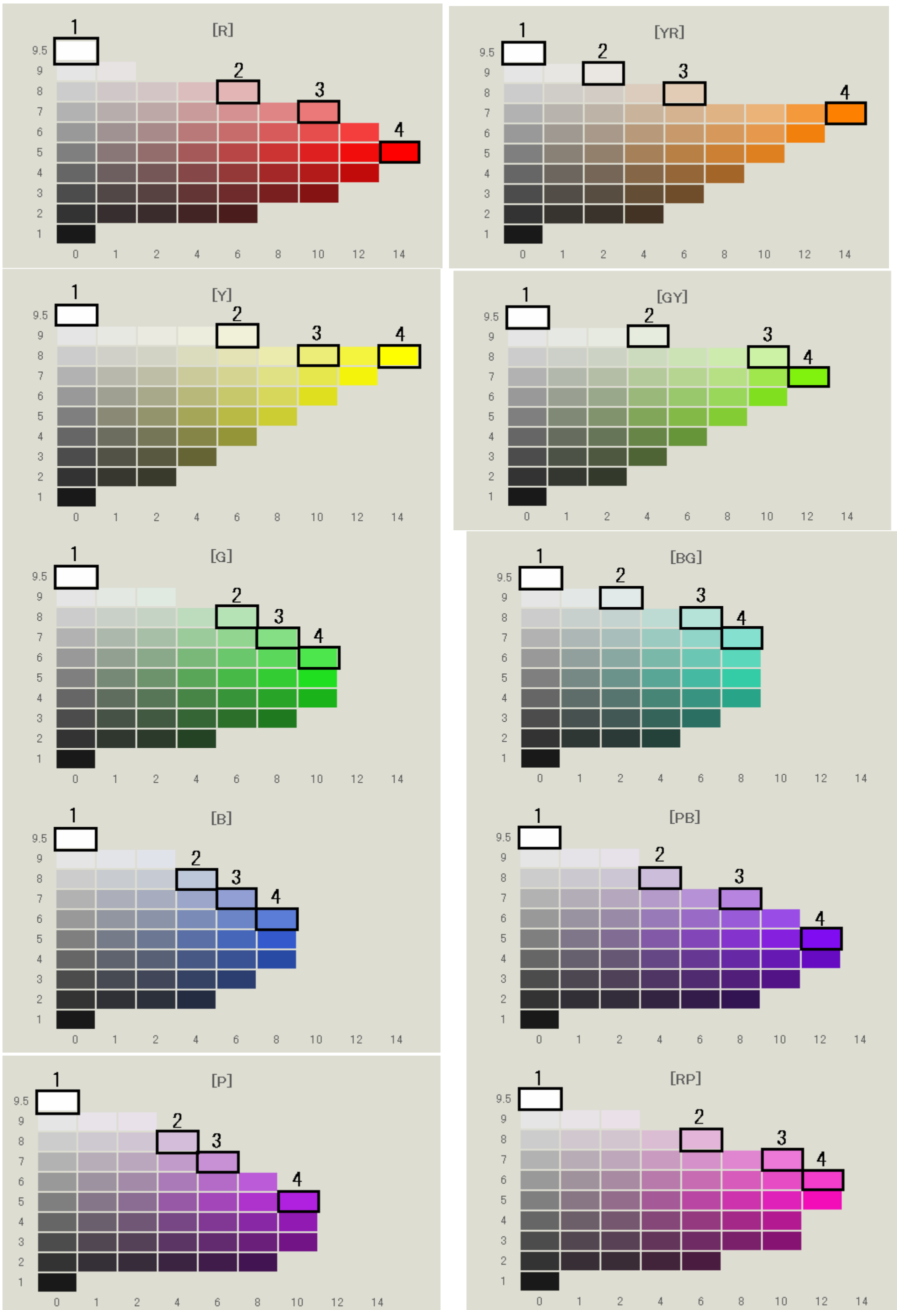


Fig. 1. Selected colors for each hue

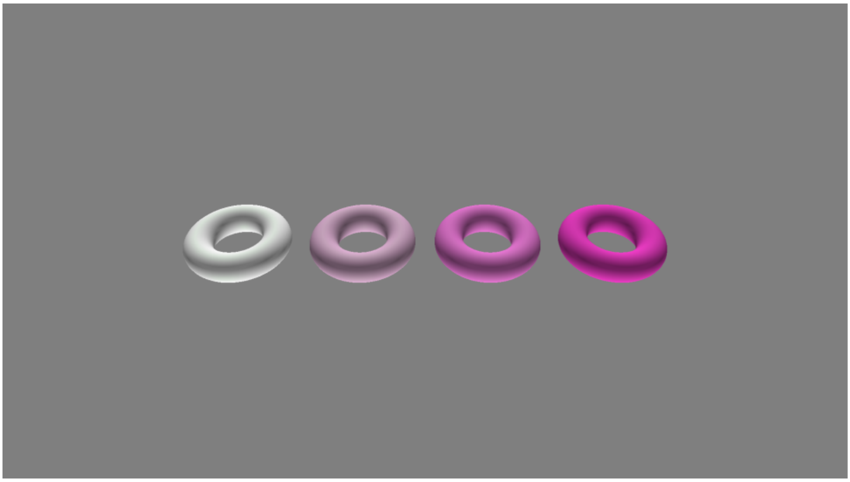


Fig. 2. Example of a set of objects

5 Experimental Results

The experiments were conducted with ten female and ten male students in their 20s with normal or normally-corrected eyesight.

The scores of the kawaii degrees were normalized on a basis of the score of white.

Fig. 3 shows an example of the results by gender. The vertical axis shows the average of the scores of kawaii degrees with the pair of brightness and saturation shown in the horizontal axis. The error bars indicate the standard deviations.

Table 1 shows the results of the analysis of variance with three elements: hue, brightness/saturation, and gender.

The following were obtained from Fig. 3:

- Colors BG#4, YR#4, and G#3 were evaluated high by males.
- Colors YR#4, RP#3, GY#3, and BG#4 were evaluated high by females.
- Color #2 of each hue was evaluated low by both genders.

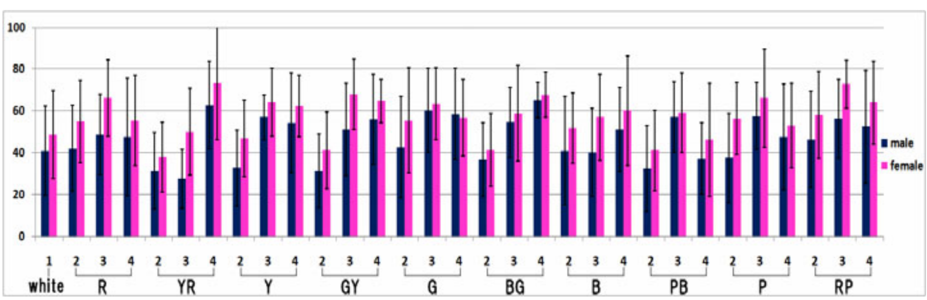


Fig. 3. Averages of kawaii degrees of each color

Table 1. Results of 3-factor analysis of variance

Factor	Sum of squared deviation	DOF	Mean square	F-value	P-value
Gender	10745.628	1	10745.628	24.168	.000 **
Hue	7949.559	9	883.284	1.987	.039 *
Brightness/Saturation	25439.626	2	12719.813	28.608	.000 **
Gender x Hue	1724.541	9	191.616	.431	.919
Gender x Brightness/Saturation	607.910	2	303.955	.684	.505
Hue x Brightness/Saturation	20509.722	18	1139.429	2.563	.000 **
Gender x Hue x Brightness/Saturation	2107.244	18	117.069	.263	.999
Error	248103.813	558	444.630		
Total	317188.043	617			

The following were obtained from Table 1:

- The main effects of gender, hue, and brightness/saturation are significant.
- The interaction effect between hue and brightness/saturation is significant.

6 Discussion

The results obtained from Fig. 3 show the following:

- Color G#3 was evaluated high by males.
- Colors RP#3 and GY#3 were evaluated high by females.
- Colors YR#4 and BG#4 were evaluated high by both genders.
- Colors with intermediate hues based on MCS got relatively high scores.

The analysis of variance results show that the scores of the kawaii degrees for colors differ by gender, hue, and brightness/saturation. The interaction effect between hue and brightness/saturation is significant, and the combinations of hue and brightness/saturation are important for evaluations of the kawaii degrees of colors.

7 Conclusion

In this study, we focused on the kawaii color of virtual objects to apply to future interactive systems and industrial products and obtained the following findings:

- The scores of the kawaii degrees for colors differ by gender, hue, and brightness.
- The combinations of hue and brightness/saturation are important for evaluations of the kawaii degrees of colors.
- Pure yellow red and pure blue green are evaluated high by both genders and are felt to be kawaii.

References

1. Belson, K., Bremner, B.: Hello Kitty: The Remarkable Story of Sanrio and the Billion Dollar Feline Phenomenon. John Wiley & Sons, Chichester (2004)
2. Murai, S., et al.: Systematic study for "kawaii" products (Third report) -Comparison of "Kawaii" between 2D and 3D-. In: Proceedings of VRSJ 2008 (2008) (in Japanese)
3. Ohkura, M., Aoto, T.: Systematic study for "kawaii" products. In: Proceedings of KEE 2007 (2007)
4. Ohkura, M., Aoto, T.: Systematic study for "kawaii" products. In: Proceedings of the Workshop on Designing Cute Interactive Media at DIS 2008, pp. 11–14 (2008), `http://www.cutemedia.org/Proceedings`
5. Ohkura, M., et al.: Systematic study for "kawaii" products (Second report) -Comparison of "Kawaii" colors and shapes-. In: Proceedings of SICE 2008 (2008)
6. Ohkura, M., et al.: Study on Kawaii Colors of Virtual Objects. In: CHI 2009, Boston, MA, USA, April 4-9 (2009)
7. Bird, S.B., Dickson, E.W.: Clinically significant changes in pain along the Visual Analog Scale. Annals of Emergency Medicine 38(6), 639–643 (2001)

Representation of Decision Making Process in Music Composition Based on Hypernetwork Model

Tetsuya Maeshiro[1], Shin-ichi Nakayama[1], and Midori Maeshiro[2]

[1] School of Library and Information Science, University of Tsukuba,
1-2 Kasuga, Tsukuba, 305-8550 Japan
{maeshiro,nakayama}@slis.tsukuba.ac.jp
[2] School of Music, Federal University of Rio de Janeiro,
Rua do Passeio, 98, Lapa, Rio de Janeiro, Brazil
mdrmaeshiro@gmail.com

Abstract. Music composition is treated as a repetitive decision making process, and represented using the hypernetwork model, the proposed model. The hypernetwork model allows more specific description of relationships among represented entities than conventional knowledge representation models such as semantic network. Music composition of a musical piece of 100 measures (performance duration of 6-8 minutes) by a professional composer is analyzed based on the description of decisions involved. Musical piece represented as musical scores are represented with hypernetwork model where the decisions are the representational units. Single or multiple decisions are related with other decisions, and quantitative similarity based on relationality among decision sequences provided by the hyperlink model enables the discrimination of various types and degrees of similarity.

Keywords: Knowledge representation, similarity relationship network, decision making, music composition.

1 Introduction

The present paper addresses the description model to represent decisions involved in composition of musical pieces. Music composition is an artistic creation process, but can also be understood as a complex sequence of decision processes that involves both formal theoretical basis of music theory that rules the decision freedom and artistic basis that expands idea creation employing mainly subjective artistic impressions and evaluations.

In the simplest view, music composition is a successive addition of single notes with specific pitch, duration and intensity. Therefore, a chord is a set of vertically aligned single notes, and a melody is a set of horizontally aligned single notes. At the finest detail level, the decision is to select a single note with determined pitch, duration and intensity. A higher level interpretation involves chord decisions on chord sequence, harmony, role distribution among different instruments, choice of measures and keys, tonality, and movements. Evidently the options and variety differ according to the type of musical piece to compose.

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 109–117, 2011.

The musical score represents the final product of composers. Intentions and thinking process of composers are hidden and not explicitly annotated, mainly because no description means exist. However, such information is crucial for musical instruments players to execute the musical piece, because better performance is possible by deeper understanding. For this reason, music courses provide analysis class which is dedicated exclusively to the analyses of musical pieces.

Even if written, composers' intentions are related solely to the final version of the music. On the other hand, more useful information is the creation history of the musical piece, described through the sequence of all decisions involved in the composition. Musical instrument players can understand deeper the “raison d'etre” of each passage, motive and aim of the composer. There is, consequently, a need for a comprehensive and powerful description model to annotate decisions.

The present model is a description model of decision processes. The following can be affirmed. (1) Descriptions of intermediate stages, progresses and background facts (reasons) of music composition processes are rare; and (2) All descriptions of the first item is written as text (natural language), thus “automatic” organization including classification and grouping is very difficult. Currently, no practical method exists to organize such information.

Conventional representation models of decision processes focalize on decision sequence and structures that arise from it. Although not explicitly stated or even noticed, these models are based on graph theory whose representation elements are nodes and links. A node represents a decision and two nodes are connected with arrow if direct temporal relation or some causal relation exists between the decisions. An arrow can be either uni- or bidirectional. More elaborated models provide multiple node types to visually differentiate decision types, such as single decision and conditional decisions, similar to computer program flow charts.

An advantage of conventional models is the easiness to grasp overall progress and structure because they offer overall view of the whole decision process, due to their simple and compact visualization.

This advantage is, however, also an disadvantage. Details of decisions can be described by adding nodes of lower hierarchy level to each node with proper annotation rules. On the other hand, details of decision transitions cannot be added, because decision transitions are simple links between nodes, and nodes cannot be connected directly to links. Furthermore, decision transitions among three or more decisions (nodes) cannot be represented since a link connects two nodes only. In reality, a decision may trigger multiple decisions, and a decision might be made subsequently to multiple decisions or integrate multiple decisions. These cases are *N*-to-*N* decisions, but impossible to be handled with conventional models that are based on graph theory.

2 Decisions in Music Compositions

A musical piece with 100 measures composed by a professional composer is analyzed to evaluate the proposed representation model of decision process. The composition process consists of four stages: (1) Creation of initial 21 measures; (2) Elaboration of the first 21 measures; (3) Creation of measures 22-49, adding to the result of the second stage; and (4) Creation of measures 50-100, the last part of this musical piece. At the end of each stage, all operations are described using composer's native language to ensure precision.

During the description, all involved decisions are classified into one of the categories based on the nature of decisions: (A) Theoretical, where the decision is derived from some music theory. Empirical (heuristics) foundations are excluded, because of the absence of theoretical basis. (B) Selective, where the composer makes an arbitrary choice from a set of candidate solutions based on chance or some theory. (C) Contingent, where the decision is arbitrary but uses composer's intuition.

It is important, however, that the decision maker (composer) is experienced, and even if the decisions are superficially apparent as random choice, the choices are based on experience accumulated by the composer. Therefore, selective choices are non-random and intuitive, which are most of the times correct or adequate [2], completely different from novices' shots in the dark.

The composition process is divided into four stages: Stage-1A, 1B, 2 and 3, following the composer's work procedure. The stages 1A and 1B refer to the first and second stages of composition of measures 1-21. The stages 2 and 3 refer respectively to compositions of measures 22-49 and 50-100, where both segments were composed in a single step. The time sequence of the composition is: Stage-1A, 1B, 2, 3. No overlap exists, i.e., no modifications in measures not belonging to the measures treated in each stage are observed.

The composed musical piece is post-tonal piece. The formal structure of such musical piece is clearly governed by the harmonic content, but not exclusively. Therefore, it is adequate to describe the decision process of musical composition, because the harmonic aspect of tonal music is completely regulated by harmonic theory, and theoretical aspect of tonal harmony is absent in atonal music.

Figure 1 is the frequencies of decision types in each composition stage. Theoretical decisions are most frequent in Stages 1A and 1B, and contingent decisions are the most frequent in 2 and 3. Small frequency of selective decisions (18.6% of all

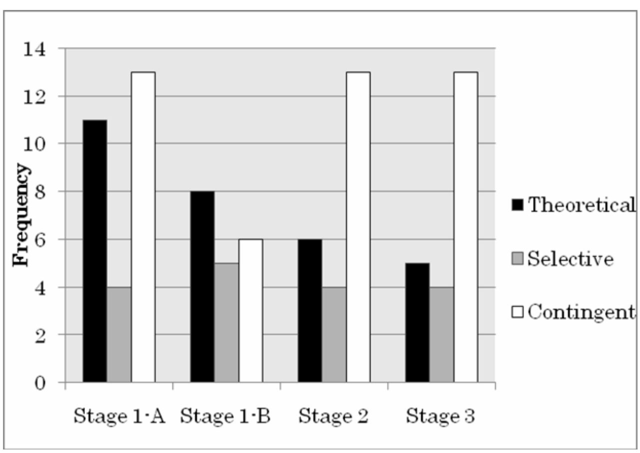


Fig. 1. Occurrences of decision types in each composition stage

decisions) suggests that professional composers create and improve musical pieces based either on theory for theoretically constrained factors or on intuition. In selective decisions, the composer has to explicitly imagine multiple solutions (options) and choose one. This is in accordance with result from decision making analysis, where an experienced professional employs the first option (solution) that comes to the mind if the solution has no apparent fault [2].

Decisions are classified into (i) framework and (ii) component decisions, based on the extent that the decisions affect. Framework decisions affect the entire musical piece, and component decisions the passsages or a part of the piece independent of the extent of influence. In other words, these are global and local decisions.

For a more detailed analysis, decisions involved until the concretization of the thematic material in the Stage S-1A are addressed in this section.

2.1 Framework Decisions

Framework decisions observed in analyzed musical piece are as follows.

1. D1A: Harmonic content is post-tonal.
2. D2A: Use pitch class set as the harmonic theme of the entire piece. The selected set is 6-Z25 in the Forte's table [1].
3. D2B: The harmonic content of the thematic material should follow strictly the selected harmonic theme.
4. D3A: Selected musical instruments are saxophone and piano.
5. D3B: Saxophone chosen as the instrument of main role to express the formal shape of the piece, at least of the part A.
6. D3C: Scheme of the piece is ternary, of the type A–B–A. The part B is dedicated to solo passages.
7. D3D: The total duration of the musical piece is between 6 and 8 minutes. This influences the tempo (bpm) and the number of measures.
8. D4C: The bpm (beats per minute) is 76.
9. D9A: Addition of the third musical instrument, the contrabass. At this point, it becomes clear that the procedure to fulfill characteristic intermittencies of main material is more effective with the existence of the third instrumental part. The new instrument not only reinforces the piano or the saxophone part, but also takes part of the metric contrast requested by the textural plane of the piece.

Other framework decisions are possible, depending on the class of musical piece, although the general characteristics are common to all musical styles. For instance, the key should be determined for tonal musical pieces, but is unnecessary for atonal musical pieces. Key change and variations are also possible. These decisions refer to the entire musical piece, and are usually executed at the beginning of the compositions. Framework decisions are, however, also executed in the middle of composition process. In the analyzed music, for instance, a third musical instrument, the contrabass, is added in the middle of the process (Decision D9A), but this decision is still in the first stage (Stage 1-A).

Fig. 2. The first version of the initial five measures. Target (affected) region of decisions are also indicated. IDs of decisions are the same used in the text, and the number indicates the sequence order of the decision.

2.2 Component Decisions

Component decisions are of varied type and account for 74% of decisions in Stage 1-A, and 92% of all decisions (Figure 1).

Component decisions are executed in steps, and each step consists of multiple decisions (microdecisions). The unit of decisions is determined by the composition task that the decision affects. A region can be the target of multiple decisions, each encompassing identical or different ranges. The affected notes or range is decision dependent, and no maximum boundary exists, possible up to a dozen of measures. The minimum boundary is obviously a single note.

Component decisions of the thematic material creation of Stage S-1A are as follows. The numbers denote the decision order, and the last alphabet indicates the order inside the decision group. For instance, D4B is the second decision of the fourth decision group.

Fig. 3. The final version of the initial five measures. For comparison purposes with the first version (Fig. 2).

1. D4A (Theoretical): For the creation of elements that constitute the thematic material, two cellular structures will be used: a fast descending movement element (element A) and pulsing element (element B), with no melodic contour.

2. D4B (Contingent): Element A contains the whole referential harmonic set, and should reach the lowest tone of the saxophone. The element B will be presented with this tone.

3. D4C (Contingent): From the movement characteristic of the element B and the rhythmic structure of elements A and B, tempo of 76 bpm is appropriate for the intended effect.

4. D5A (Contingent): The element A should sound starting with an explosive intensity, then followed by energy dissipation, represented with initial accent with diminuendo.

5. D5B (Contingent): The element B is characterized by higher intensity and articulations of counterpoint, indicated with accent and staccato, valuing the pause that is important to realize the theme.

6. D6A (Theoretical): The initial phrase is complete by joining to the principal motif (A+B) a complementary material by amplification derived from the element B.

7. D7A (Theoretical): A second phrase structure to establish a new symmetry is necessary to achieve the first formal closing point, on the level of complete phrase structure.

8. D7B (Contingent): The second phrase has the ornamentation type identical to the first phrase, keeping the economy and homogeneity of harmonic resources.

9. D7C (Contingent): The element A is modified with amplified ornamental movement, followed by a new element B without the central articulation.

10. D7D (Contingent): The second phrase is terminated with the complementary material of the initial phrase, but with reconfigured metric structure for fitting reason.

11. D8A (Selective): The imprint of the piano part of this section is the filling of intermittencies that exist in the main part's contour executed by the saxophone.

3 Hypernetwork Model

The model to represent the decision making process, the hypernetwork model [3], is extended from the bipartite representation of the hypergraph [4]. The hypergraph model, on the other hand, has more representation capability than conventional knowledge representation models that are based on graph [6], such as semantic network [5], frame, and ER-model [7]. Conventionally used decision sequence representation is also a graph. Basically, the hypernetwork model follows basic definitions of semantic networks, where a node is connected to other nodes (1) to specify the nodes or (2) when nodes are related by some relationship.

A uniqueness of the hypernetwork model is the existence of three types of description elements, equivalent to the types of nodes. Graph and hypergraph models consist of nodes and links connecting the nodes. In decision sequence representation, a node represents a decision, and a link connects two or more decisions in sequence relationship. A link of the graph model can connect only two decisions, and a link of the hypergraph (hyperlink) connects any number of decisions. The bipartite representation converts the links into nodes, denoted relation nodes, hence two types of node exist: the vertex node and the relation node. The vertex node serves to represent decisions (entities), and the relation node to describe relationships among decisions. An analysis of knowledge property, however, indicates that a third type of node is necessary, the attribute node, to specify the properties of vertex nodes and relation nodes. Therefore, conventional representation models present at least two flaws: (1) relationships among multiple entities cannot be represented, and (2) representation is incomplete since attributes are not provided. The hypernetwork model resolves both problems.

In the context of decision representation, details or properties of a concept represented by a vertex node can be specified in two ways: by attachment of attribute nodes, or by relating to other vertex nodes through relation nodes. Combination of the two descriptions is also possible. The attribute node exists to specify any of three node types. Table 1 indicates the connectivity constraints among three node types. Two connections are prohibited: between vertex node and vertex node, and between relation node and relation node, constraint imposed from their role in hypergraph. Table 1 is symmetrical on diagonal axis although the directionality of links depends on the context and what the network represents.

Table 1. Connectivity among vertex node, relation node and attribute node

	Vertex node	Relation node	Attribute node
Vertex node	–	Connect	*Connect*
Relation node	Connect	–	*Connect*
Attribute node	*Connect*	*Connect*	*Connect*

4 Discussions

The proposed model focuses on the process of composition, and not on methods to represent the final outcome.

Uniqueness of the presented analysis is that decisions are of intermediate stages, related to the fact that composition process of musical pieces is a succession of intermediate versions. Surely some musical pieces were created in a single stage, but most are succession of intermittent generation and improvement. It is known that some large musical pieces such as symphony have taken several years to be completed.

The model is also useful for musical piece analysis by the performers, such as instrument players and conductors, to understand the “raison d’etre” of individual passages and of the whole piece. This is important, because better execution is possible by deeper understanding.

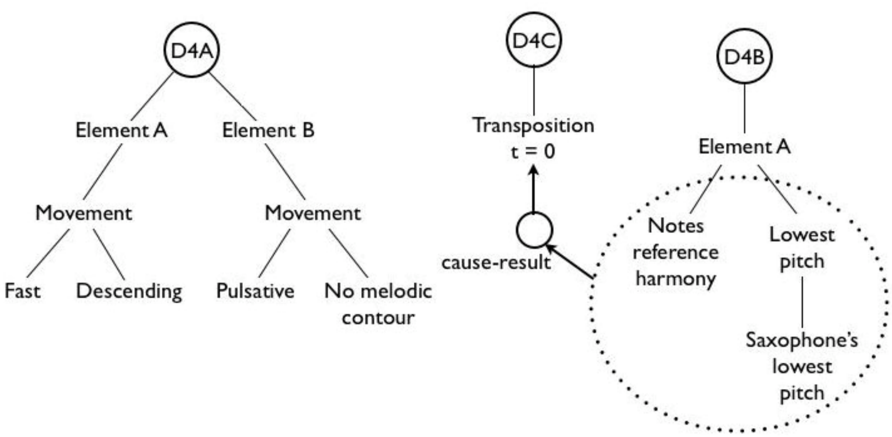


Fig. 4. Hypernetwork representation of the decisions D4A, D4B and D4C. “Element A” in D4A and D4B can also be related by identity relationship. In D4B, a higher order hierarchical level is created to related with D4C.

An application of the hypernetwork model is the evaluation of decision processes. Quantitative comparison of decision process based on the structure of represented hypernetwork allows the design of composition strategy. The quality of the whole decision process is measured by the quality of composed music. Obviously the quality of the generated music is affected by many factors unrelated with the decisions involved in the composition. However, the structural analysis can be a useful measure.

Attribute node linked to a vertex node specifies or defines the properties of the entity represented by the vertex node. Attribute node linked to another attribute node is homologous to the previous case, and it defines the properties of the quality or concept expressed by the other attribute node. On the other hand, attribute node connected to a link node is the element absent in conventional representation models. This connection enables a detailed specification of relationship among vertex nodes. Note that the relationship treated here is N-ary which covers the binary relationship, the only relationship that conventional models can represent.

The ability to assign attributes to relationship is essential to qualify and then quantify the similarity relationships that are simply labeled "similar" in semantic networks, for example. Furthermore, attribute nodes connected to link node are specified with more details by further connecting attribute nodes, generating a multi-level hierarchical structure of attributes. Any type of relationship is specified in same manner, but this paper focus on "similarity", a very broad conception, meaning anything between identical and different. Since vertex nodes and relation nodes can connect to attribute nodes, relation nodes are related to other relation nodes. Consequently, a decision network with higher density than conventional models emerges.

Knowledge description using representation models is difficult, and no definitive model exists. The proposed representation model is useful to embed into knowledge base system to search similar composition decisions/operations/techniques, mainly for professional composers. Once composition stages and types of musical piece are specified, it is useful for composers to widen their views and overview multiple composition solutions as their creative process can be enriched by choosing one of solutions, combining them, or improving them.

References

1. Forte, A.: The Structure of Atonal Music. Yale Univ. Press, New Haven (1977)
2. Klein, G.: Sources of Power: How People Make Decisions. MIT Press, Cambridge (1999)
3. Maeshiro, T., Maeshiro, M., Shimohara, K., Nakayama, S.: Hypernetwork model to represent similarity details applied to musical instrument performance. In: 13th International Conference on Human-Computer Interaction, pp. 866–873 (2009)
4. Berge, C.: Hypergraphs: Combinatorics of Finite Sets. North-Holland, Amsterdam (1989)
5. Quillian, M.R.: Word concepts: a theory and simulation of some basic semantic capabilities. Behavioral Science 12(5), 410–430 (1967)
6. Berge, C.: The Theory of Graphs. Dover, New York (2001)
7. Date, C.J.: An Introduction to Database Systems, 8th edn. Addison-Wesley, Reading (2003)

Some Issues toward Creating Human-Centric Services

Hirohiko Mori

Department of Industrial and Management Systems Engineering,
Tokyo City University
1-28-1, Tamazutsumi, Setagaya, Tokyo, 158-8557, Japan
hmori@tcu.ac.jp

Abstract. Some issues about creating some human-centric services and contents are described and discussed, especially in ubiquitous computing environments. The term of "ubiquitous" are used in the different meanings in different situations. So, I, first, try to identify each meaning and, according to this arrangement, I discuss some issues how to create human-centric services in the academic ways.

1 Introduction

It has been a while since the term "Our age is a ubiquitous computing society" began to be used generally; however, what "ubiquitous" really means is seldom understood. In fact, the term *ubiquitous* is used in connection with a variety of different uses such as "ubiquitous society," "ubiquitous network," or "ubiquitous computing," and each term has a different meaning. So in this paper, I shall first define a ubiquitous society in my own way and consider the point of view that the services offered in society should be produced.

2 What Is a Ubiquitous Society?

The term *ubiquitous* refers to a "ubiquitous network," "ubiquitous society," or "ubiquitous computing" and began penetrating into the general public, I suppose, as the more popular "ubiquitous network." At first, the term *ubiquitous network* meant, for example, an environment where one could connect to the Internet anytime and anywhere, such as in front of convenience stores, in train yards, on trains, or in towns, and in many situations, it meant Wi-Fi networks available everywhere. Thereafter, the term began to include sharing information and an environment where every single device was connected to a wired network in the home.

A ubiquitous society means a society where services and information are offered based on a ubiquitous network infrastructure. The environment where one can simply connect to the Internet anywhere, for example, with a more extensive capability of searching for information from a browser, as in, for instance, the POP-type information supplement that suits the situation: information is displayed on electronic bulletin boards in stations or as one leaves the station ticket gate, surrounding area information

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 118–121, 2011.

is displayed on a mobile phone. In these instances location information is simply used; however, the goal is to realize a variety of situations as in what a user intends to do and is interested in and to offer services suitable for that.

Also, a variety of information is currently shared through DLNA-enabled devices or audiovisual equipment. The situation is becoming ubiquitous where videos recorded onto a hard disc recorder in a living room can be played on a game machine in another room. I suppose that this is the beginning of a ubiquitous society. This is merely simple information sharing for now; however, it essentially means that different kinds of machines will coordinate with each other. That is, in the next stage, computers embedded in home appliances will begin working together via information-communication technology and switching to produce some sort of service. For instance, for digital home appliances, instead of videocassette recorders incorporating a clock function, the recording function of a VCR programmed by an alarm clock and video recorder information would be communicated via a home network.

In addition, beyond electronic devices, these networks will expand into common personal items, such as furniture or accessories. Even now, personal items and non-electronic devices have information and communication functions from IC tags, such as on vegetables in order to grasp producer information. That is, everything has an information and communication function in a ubiquitous society, where they will be producing coordinated services.

From the above, when thinking of services according to the term *ubiquitous*, the following can be assumed:

(i) Information and communication devices are embedded onto one's surrounding environment and personal items.
(ii) The devices have communication functions and can create and coordinate services.
(iii) They recognize a variety of situations and offer suitable information. While the term *ubiquitous* meant "anywhere" in a ubiquitous network, the concept "to anything" and "for each person depending on the circumstances" are part of a ubiquitous society.

3 Service Creation in a Ubiquitous Society

When thinking of the service layers in a ubiquitous society, services or information offered should be based on human-centered design because a ubiquitous system must fit different lifestyles.

Service layers will extract user needs based on an analysis of the requirements in general human- centered design. However, it is difficult to apply these methodologies to a ubiquitous era. In these methodologies, it is essential to extract user needs through interviews.

An analysis of requirements identifies mostly dissatisfaction with the current system, and new services not in existence today.

As a new attempt to resolve these issues, Go, et. al. suggested the photo scenario method. Unlike traditional marketing research, this method extracts potential needs from a small volume of data, clipping part of life with a photo diary, and considers whether it is possible to embed the needs.

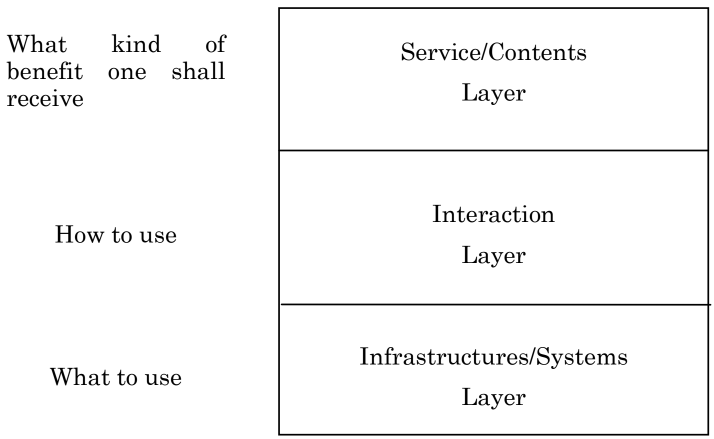


Fig. 1. Ubiquitous Three Layers model

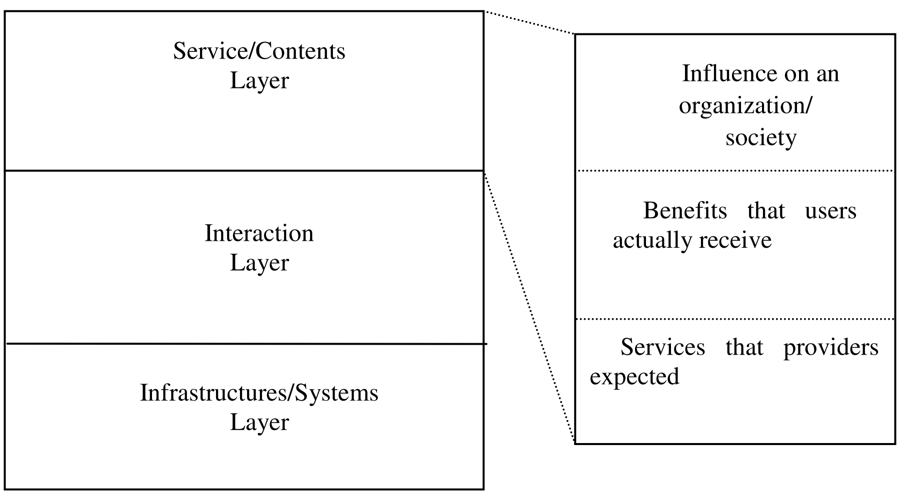


Fig. 2. Detailed Service Layers

My opinion on having attempted this method so far is that what is important is not to think imprecisely about whether discovered needs can be clipped as photos in a diary, but to try to create adaptive services and to consider the effectiveness. Also, rewriting a service scenario and an interaction scenario more than once when creating a use scenario for the services is effective. That is, design separately what services users receive, what system is used, and how it is used. This makes it possible to clarify the purpose of the services, to think of different infrastructures or interaction so as to achieve the services, and to consider how infrastructure or interaction influences the services.

Finally, although Mori suggests considering three levels (Figure 1), i.e., the levels of system infrastructure, interaction, and service contents, separately when thinking of services in a ubiquitous society, I think the service levels should be subdivided further (Figure 2). That is, the goal of the service developer, the use of the services, and how the services influence a society and organization. Users do not always use services as the developer intended. Although it will be necessary to clarify that point, this is difficult at present and shall be a future issue.

References

Mori, H.: Ubiquitous three-layer model and human centered design. To achieve a human-centered ubiquitous society-, measurement and control regulation 47 (2008)

Go, K., et al.: Supporting Omoiyari: A Design Issue in Post-Ubiquitous Society. In: SICE-ICCAS 2006, CD-ROM Proceedings (2006)

A User-Centric Metadata Creation Tool for Preserving the Nation's Ecological Data

Fatma Nasoz[1], Renee C. Bryce[2], Craig J. Palmer[3], and David J. Rugg[4]

[1] The Lincy Institute, University of Nevada, Las Vegas, 4505 S. Maryland Parkway, Las Vegas, NV 89154-3067, USA
[2] Computer Science Department, Utah State University, Old Main 414, 4205 Old Main Hill Logan, UT 84322-4205, USA
[3] The Harry Reid Center for Environmental Studies, University of Nevada, Las Vegas, 4505 S. Maryland Parkway, Las Vegas, NV 89154-4009, USA
[4] U.S. Forest Service, Research and Development, 1 Gifford Pinchot Drive Madison WI 53726-2366, USA

fatma.nasoz@unlv.edu, renee.bryce@usu.edu, craig.palmer@unlv.edu, drugg@fs.fed.us

Abstract. This paper describes the methodology to develop a metadata editing environment that (1) is scientist-friendly, (2) promotes accurate recording of metadata, and (3) enables interoperability across multiple metadata standards. Scientist-friendly software will increase scientists' willingness to create metadata and the quality of that metadata. This study identifies the usability issues in an existing metadata creation tool and initiates the development of a state-of-the-art user interface that allows the biological research community to more accurately record and share data.

Keywords: Metadata, user -centric, usability, heuristics.

1 Introduction

Current metadata tools suffer from several usability issues that result in users often under-recording and incorrectly recording data using consistent metadata standards. This project's objectives are to identify usability issues with existing metadata creation tools and use this information to develop a metadata tool that is scientist-friendly and promotes accurate recording of data. This scientist-friendly software will improve the quality of that metadata, which in turn will improve the frequency in which scientists are willing to create metadata. Improved interoperability will make it easier for biologists outside of the originating discipline to understand and search the data being shared. It can also aid re-use within a discipline when multiple metadata standards are used. The project's primary product will be an open-source metadata editor package that is based on several iterations of usability feedback from the research community.

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 122–131, 2011.

2 Motivation

2.1 Increased Attention to Long-Term Data Management

Biological data are collected in different locations, at different temporal and spatial scales, and using diverse methods. In most cases, however, these data are collected and analyzed digitally. The re-use of such digital data for various scientific purposes has attracted attention [1, 2]. High quality metadata are critical aids to the re-use of data, whether by the collecting scientist or by data users outside of the collection team [3] and they are part of any well-designed data management plan.

Metadata describe the *who, what, when, where* and *how* about data [4]. The quality of metadata records determine the ease of discovery, acquisition, integration, comprehension, and preservation of data. Over time, low quality or missing metadata records result in losses in information content for a given data set – an increase in "data entropy" [5,6], and make re-use of the data set difficult or impossible.

There are numerous barriers to quality long-term data management. Cultural issues in the research community have been a serious barrier, but two developments hold promise for reducing this barrier. First, an increasing number of biology-focused scientific journals are requiring that research data supporting published articles be archived and shared. Second, the National Science Foundation announced the requirement of data management plans for all grant proposals by the end of 2010 [7]. Other barriers are more technical than cultural. One such barrier is the difficulty many scientists experience when attempting to use metadata authoring software [8]. These difficulties result in scientists investing the effort to create metadata less frequently and generally creating lower quality metadata than they or the data users would like. Another technical barrier is deciding which metadata standard to use. Disparate standards make writing metadata harder for scientists and make using metadata to understand research data more difficult for data consumers. This situation is a barrier both within and across disciplines.

2.2 Challenges Faced in Writing Metadata Documents

A study [9] of metadata authoring identifies numerous challenges that authors face when writing compliant metadata documents. These challenges result in frustration, which leads to non-adoption of the process to create metadata or incomplete and inconsistent creation of metadata. Identified challenges include:

- Most researchers write metadata records infrequently, perhaps just once or twice a year.
- The majority of biologists (and others) who attended metadata training workshops feel they do not have an adequate command of existing metadata tools, and characterize those tools as too "complex" and "cumbersome".
- Certain elements of available metadata standards are defined differently by different users, and on occasion, by the same user at different metadata writing sessions. Study participants called the element definitions "vague".
- Researchers responded favorably to an interview format for creating metadata records, but this feature is not commonly available in existing tools with the exception of the Ecological Metadata Language (EML) tool. When deciding on

keywords for metadata records, authors select their own terms rather than use those from an established controlled vocabulary.
- Creating metadata is time consuming and costly.
- Potential metadata authors seek training, but training workshops are limited in size, frequency, duration, and location.
- Given the same dataset, two metadata authors often produce significantly different documents.

When we performed usability testing on existing metadata creation tools before preparing this proposal, we found similar results to those reported in [9].

3 Methodology

The starting point of our project was Metavist 2005 (v1.6) [10] (Fig. 1), widely used open source software developed by one of the authors of this paper. Since its development, Metavist 1.6 has been used by a variety of users to enter and access metadata.

We initiated our study by performing a usability evaluation of Metavist 1.6 using Nielsen's set of heuristics [10]: visibility of system status; match between system and

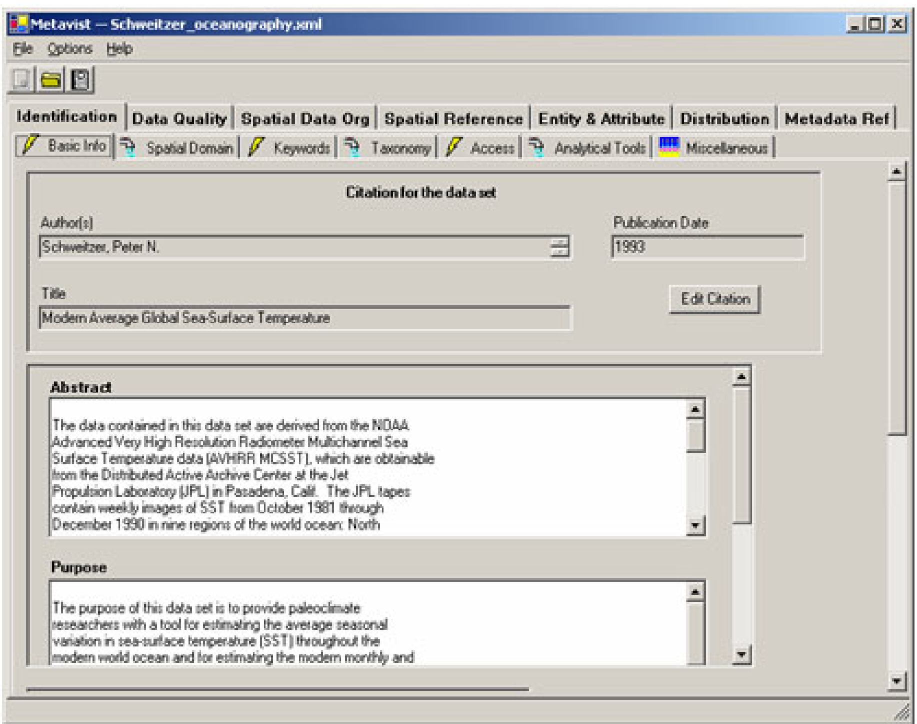


Fig. 1. Metavist 2005 (v1.6)

the real world; user control and freedom; consistency and standards; error prevention; recognition rather than recall; flexibility and efficiency of use; aesthetic and minimalist design; help users recognize, diagnose, and recover from errors; and help and documentation. Our usability the study identified the subset of heuristics with which Metavist v1.6 did not conform.

Match Between the System and the Real World. "The system should speak the users' language, with words, phrases and concepts familiar to the user, rather than system-oriented terms. Follow real-world conventions, making information appear in a natural and logical order." [11]. In Metavist 1.6, the several icons that represented various functions and fields were confusing as they did not appropriately match the intent/meaning of those functions and fields. As seen in Fig. 2, the notebook icon represents the "save" function but the image did not reflect the real world object where a regular floppy disk image would be more appropriate to represent the action. Similarly, the "open" operation had a mismatching icon. Also, the "Mandatory", "Mandatory If Applicable", and "Optional" sections were represented with lightning, running tab, and birthday cake respectively. These icons did not reflect the meaning behind the sections.

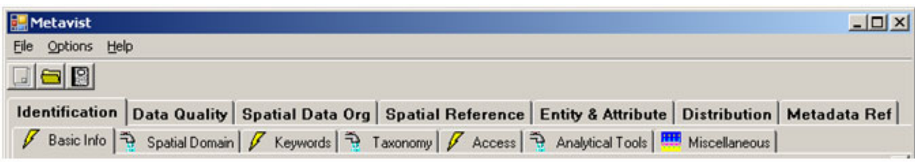


Fig. 2. Metavist 1.6 – Icons

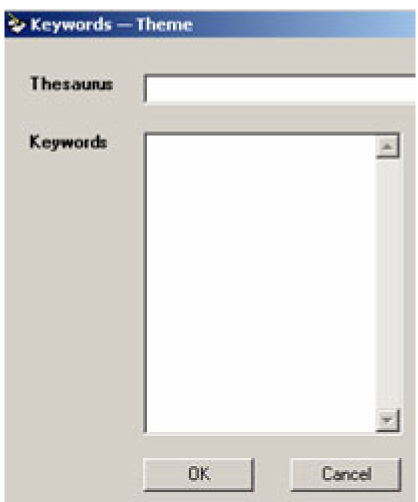


Fig. 3. Metavist 1.6 – Keywords Dialog Box

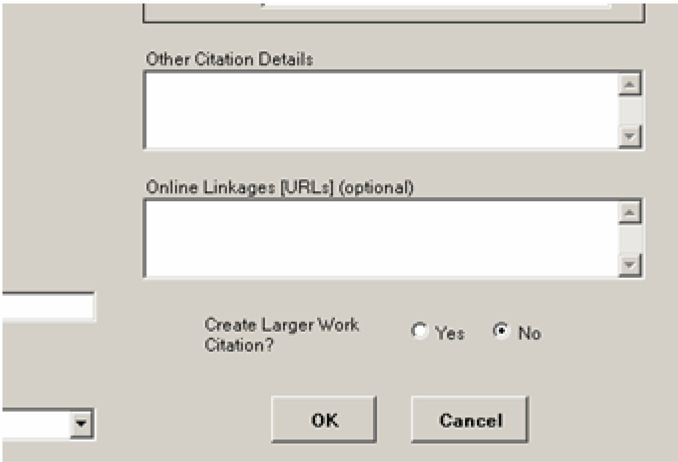


Fig. 4. Metavist 1.6 – Citation Dialog Box

Consistency and Standards. "Users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform conventions." [11]. Syntaxes, icons, layout, and functions used in the system should always be consistent. Metavist 1.6 had several buttons with different width and lengths, or with different font styles. In two dialog screens (Fig. 3 and Fig. 4) the "OK" and "Cancel" buttons' locations and types were different from each other's. Another usability issue with the dialog boxes was that they appeared on different locations of the screen each time, so it was hard to keep track of the location of a new dialog box screen.

Error Prevention. "Even better than good error messages is a careful design which prevents a problem from occurring in the first place. Either eliminate error-prone conditions or check for them and present users with a confirmation option before they commit to the action." [11].

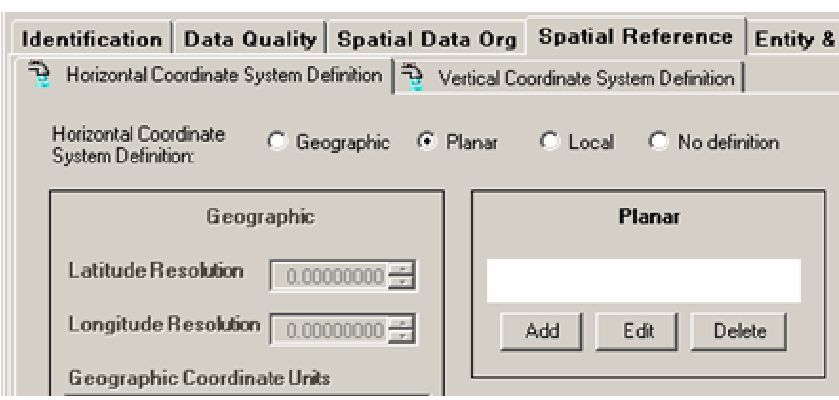


Fig. 5. Metavist 1.6 Horizontal Coordinate System Definition Section

Fig. 5 shows that the "Add", "Edit", and "Delete" buttons were enabled in the "Planar" section although there was no item available in the list box or no item was selected. If user tried to use one of those buttons, a warning dialog would appear. This usability issue can easily be fixed by disabling the buttons until the user adds or selects an item in the list box.

Similarly Fig. 6 shows the case where the "OK" button on the Standard Order Process Dialog was enabled even though no format was selected. If the "OK" button was to be clicked without choosing a format an error dialog box appears Fig. 7).

In this case, the "OK" button should remain disabled until the mandatory fields are completed by the user, so the user would not face another dialog box to worry about and would understand the requirements of the metadata creation for specific sections of the standards.

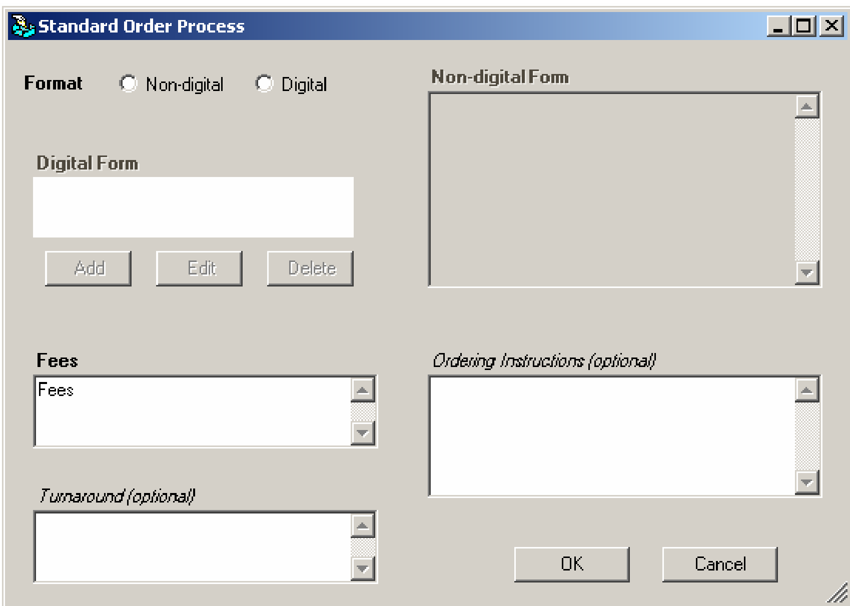


Fig. 6. Metavist 1.6 Standard Order Process Dialog

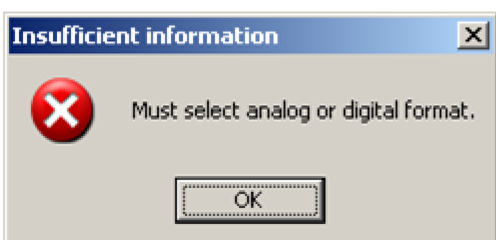


Fig. 7. Metavist 1.6 Standard Order Process Dialog Error Message

3.1 Ongoing Work

As described in the previous section, heuristics were used to evaluate the usability of Metavist 1.6. Also, usability issues pointed out by the users of existing metadata creation tools were taken into account while implementing the improvements. Existing functionalities of Metavist 1.6 were kept the same; instead we aimed to improve the usability of those functionalities.

An iterative and incremental design approach was adopted for the software development lifecycle. In the requirements analysis phase, existing metadata tools were analyzed briefly to identify usability issues and needs and comments of existing users were gathered.

Once we identified and prioritized the modifications to the software and created the initial design, the implementation phase started by the converting the existing framework version into .NET framework 3.5. The system was kept as a single-user desktop application as it was before. Implementation was done on Microsoft .NET Framework 3.5 along with Visual Basic programming language.

Most of the improvements were made on the Entity & Attributes section of Metavist 1.6. New features that were added as part of the improvement are:

Auto Save. One of the main functions that the application was lacking was an auto save or confirmation before exit. Auto save and auto recovery features were added so that every time a change is made, the file is saved either as a temporary file (if the file is not saved yet) or as the actual file itself. Therefore, unexpected computer problems or application problems will not lead the loss of the metadata entry made up to that point.

Also, with the modification, in case the user wants to close the application, open a new metadata XML file, or create a new one, they are asked to confirm whether they want to save the changes before exiting the current metadata set as shown in Fig. 8. If no changes have been made since the last save operation, the confirmation will not appear.

Improved Navigation. Navigation was modified so that the user can navigate through the section with just one click. This is shown in Fig. 9. The user can directly access the desired section from the left panel where the navigation tree view is located. Also they can navigate through the sub-sections from the main panel. The users are also free to hide the navigation pane anytime.

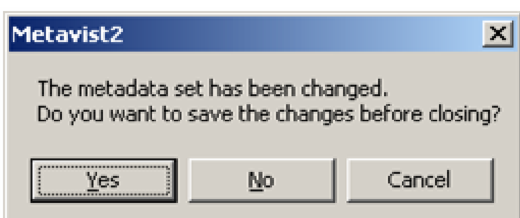


Fig. 8. Confirmation Dialog on Exit

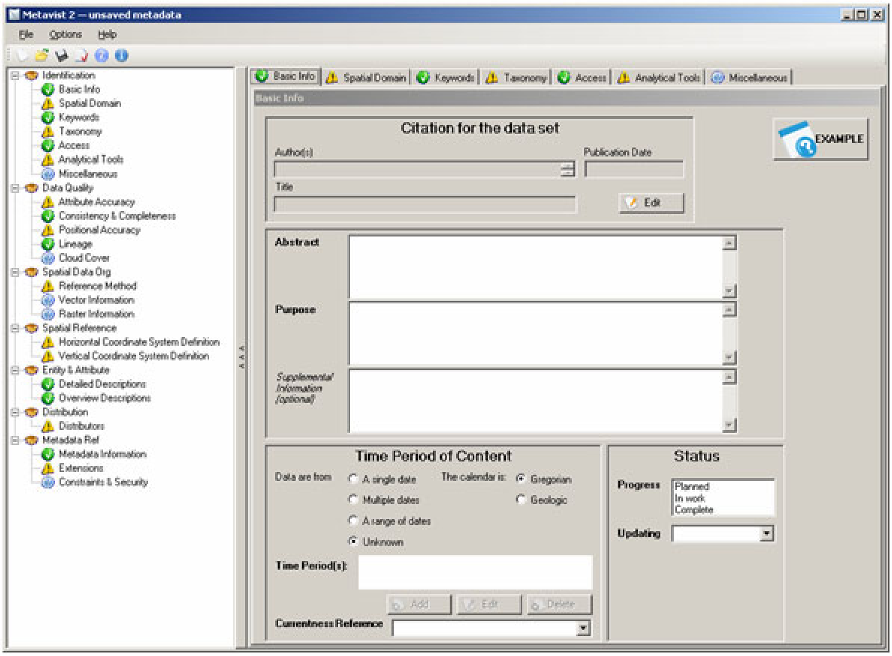


Fig. 9. Metavist 2.0 – Basic Info Section

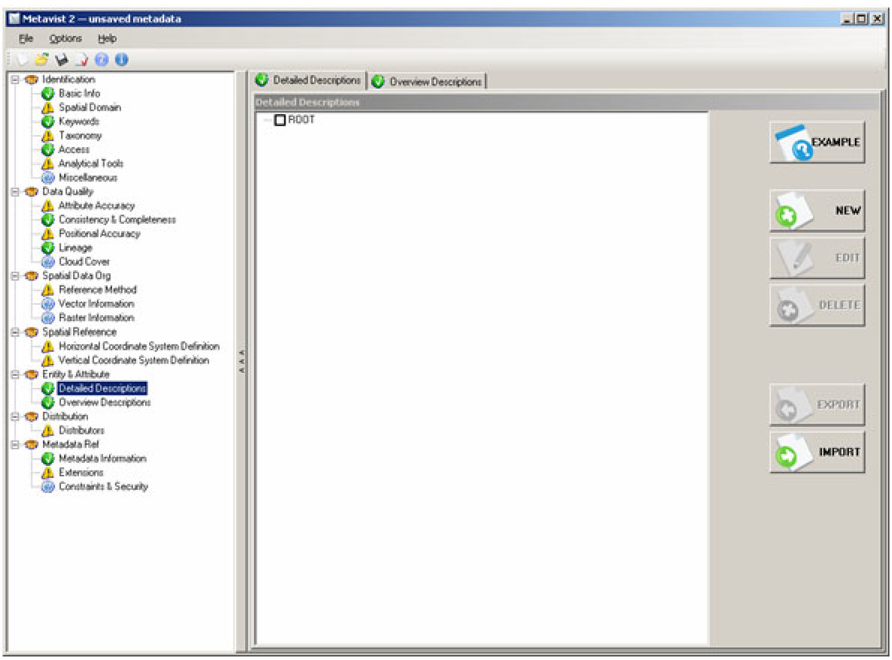


Fig. 10. Metavist 2.0 – Entity and Attributes Section

Once the beta version was created, potential users were given the software to test and provide feedback to us regarding the usability of it. Based on user feedback final adjustments were made. Fig. 9 and Fig. 10 show snapshots from the current state of Metavist 2.0.

4 Future Work

One of our future goals is to develop our software for the open source domain so that others may not only use the tool as-is, but may also build upon it when future metadata standards are released. Indeed, a major issue with current metadata tools is that when a new metadata standard is released, tool developers create new releases for the new standard, but they often no longer support backwards-compatibility within the new release. Therefore, we plan to use XML to make it simple to maintain compatibility with multiple versions of standards so that users have continuity in their ability to record new metadata and also to search older records. Indeed, the metadata standards selected for this project already have some features built-in (XML schema, XML style sheets, etc.) that the team can use. The Biological Data Profile and Ecological Metadata Language standards each have open source software that supports the creation of compliant metadata. We will utilize these two packages in our implementation of the cross-standard abstract metadata framework. Upon the release of version 3 of the Biological Data Profile by National Biological Information Infrastructure (NBII), we will create an XML schema for use in our own tool, but we will also release it to the public domain for use in any project.

The open source applications will include both a desktop application for organizations that have security measures in place that prevent internet access and a web application for organization that want to share data. Both the desktop application and the web-based application will be implemented using Microsoft Visual Studio 2010 Integrated Development Environment (.NET Framework 4). This environment will allow the team to develop and test the software collaboratively using the Team Foundation Server. The application will run on a variety of platforms, including multiple of versions of Microsoft Windows and Linux (supported by the Mono Project). For web-based applications, cross-platform use will be accessible through a web browser capable of handling a rich internet application (e.g., an AJAX-based application).

References

1. National Science Board Committee on Programs and Plans. NSB-05-40, Long-lived digital data collections: enabling research and education in the 21st century. National Science Foundation (2005)
2. US Interagency Working Group on Digital Data. Harnessing the power of digital data for science and society. Report of the Interagency Working Group on Digital Data to the Committee on Science of the National Science and Technology Council (2009)
3. Porter, J.H.: A brief history of data sharing in the U. S. Long Term Ecological Research Network. Bulletin of the Ecological Society of America 91, 14–20 (2010)

4. Brunt, J.W.: Data management principles, implementation and administration. In: Michener, W.K., Brunt, J.W. (eds.) Ecological Data: Design, Management and Processing, pp. 25–47. Blackwell Science, Malden (2000)
5. Michener, W.K., Brunt, J.W., Helly, J.J., Kirchner, T.B., Stafford, S.G.: Nongeospatial metadata for the ecological sciences. Ecological Applications 7, 330–342 (1997)
6. Michener, W.: Meta-information concepts for ecological data management. Ecological Informatics 1, 3–7 (2006)
7. National Science Foundation. Scientists seeking NSF funding will soon be required to submit data management plans. Press Release 10-077 (2010)
8. Crystal, A., Greenberg, J.: Usability of a metadata creation application for resource authors. Library & Information Science Research 27, 177–189 (2005)
9. Landis, E. Enabling FGDC Metadata Development in Biodiversity and Ecosystem Sciences (2005), `http://www.fgdc.gov/library/whitepapersreports/sponsored-reports/Enabling_Metadata_Development.pdf`
10. Rugg, D.J.: Creating FGDC and NBII metadata with Metavist 2005. Gen. Tech. Rep. NC-255. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Research Station (2004)
11. Nielsen, J.: Heuristic evaluation. In: Nielsen, J., Mack, R.L. (eds.) Usability Inspection Methods. John Wiley & Sons, New York (1994)

Comparison between Mathematical Complexity and Human Feeling

Masashi Okubo[1], Akiya Togo[2], and Shogo Takahashi[1]

[1] Doshisha University, 1-3 Miyakodani, Tatara, Kyotanabe, Kyoto,610-0321, Japan
[2] Nihon Unisys Co. Ltd. 1-1-1 Toyosu, Koto-ku, Tokyo, 135-0061, Japan
mokubo@mail.doshisha.ac.jp, akiya.togo@unisys.co.jp, dtk0773@mail4.doshisha.ac.jp

Abstract. Recently, we have often the opportunity to shop for something on the computer display connected to the Internet. However, it is often said that the product arrived at the home is something different from it presented on the computer display. We have studied into the differences between the shape evaluation in virtual space and in real space. And it is indicated that the shape evaluation in virtual space is slightly difference from it in real space, especially, in case of evaluation of product which has the complex contour[1]. In this research, we are focusing on the complexity of contour. The complexity is defined by mathematical methods, for example, Hurst exponent in fractal geometry. In this paper, we have proposed the algorithm which makes the curve lines based on the Hurst exponent. And using these curve lines made by proposed method, we have investigated that the mathematical complexity is equal to the complexity which people feel by sensory evaluation or not. The result shows that Hurst exponent almost can show the complexity which people feel, however, it is difficult in case of the cyclic curves.

Keywords: Curve complexity, Sensory evaluation, Bradley-Terry model.

1 Introduction

A complexity of curve is defined by mathematical method absolutely. Hurst exponent in fractal geometry can indicate the complexity as a number which is between 0 and 1. On the other hand, the complexity of curve by human feeling cannot be defined absolutely, but relatively. In this paper, we try to compare the human feeling complexity to mathematical complexity by sensory evaluation and statistical analysis. And we have tried to define the human feeling complexity as an absolute number.

First of all, we propose the curve generation method based on Hurst exponent. Then, some curves were made by the proposed method. Using these curves, the sensory evaluations by paired comparison were performed. To compare the mathematical complexity to human feeling complexity, we applied the Bradley-Terry model to the result of paired comparison. The model enables to determine the complexity based on

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 132–141, 2011.

the paired comparison. Finally, we tried to define the human feeling complexity as an absolute number.

2 Curve Generating Algorithm Based on Hurst Exponent

2.1 Hurst Exponent

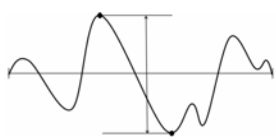
Complexity is able to be defined by Hurst exponent from mathematical viewpoint. Hurst exponent can be obtained by some methods. In this study, the scale transformation analysis has been used to define the complexity. Hurst exponent (*He*) is shown by the expression as follows:

$$He=m-D \tag{1}$$

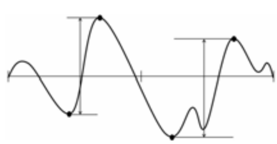
Here, m shows the Dimension [m =2 in this research], D shows the fractal dimension [$1<D<2$]. Generally, it is said that an object which has the small Hurst exponent is complex.

The scale transformation analysis pays attention to the maximum fluctuation of the curved line. The curved line divided into 2^{n-1} parts (See. Fig.1), and the maximum fluctuation of each part is measured from minimum value and maximum value of the part. The average of maximum fluctuation of 2^{n-1} parts and n is plotted on the log-log graph, and the slope of the regression line is the Hurst exponent.

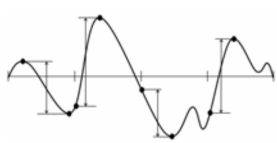
For example, to obtain the Hurst exponent of the curve shown in Fig.2, Table 1 shows the averages of fluctuations between the min. and max. values. Fig.3 shows the



(a) First step



(b) In case of dividing into 2 parts



(c) In case of dividing into 4 parts

Fig. 1. How to obtain Hurst exponents

Fig. 2. Example of curve

Table 1. Parameters of curve in Fig.2

Partitions	Reciprocal number of partitions(log10)	Average of difference between max. and min. value(log10)
1 part	0	2
2 parts	-0.301	1.909716
4 parts	-0.602	1.819412
8 parts	-0.903	1.729094
16 parts	-1.204	1.638795
32 parts	-1.505	1.548497
64 parts	-1.806	1.458196

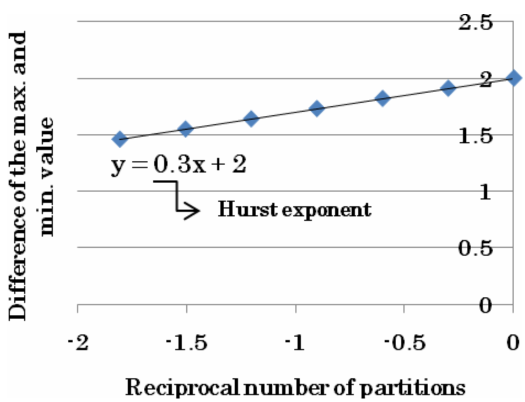


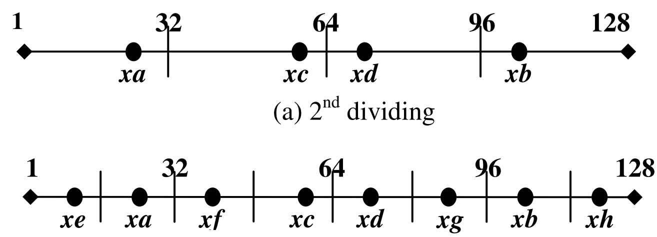
Fig. 3. How to obtain Hurst exponent

log-log graph in which the data from Table 1 are plotted on. Then, the slope of the regression line is the Hurst exponent.

2.2 Curve Generation Algorithm

We are proposing the method to generate a curve based on the Hurst exponent. The procedure for generating the curve $f(x)$ which is consisting of 128 points is as follows.

(i) Firstly, 2 points from 128 are chosen randomly. Then, one point *xa* is chosen from 1-64 and the other one *xb* is done from [65-128]. These will be the points that indicate global maximum and minimum values of the curve. Secondly, dividing the 128 points into two parts, one is consisting of 1-64 and the other is consisting of 65-128. If $f(xa)$ is global maximum value and *xa* is smaller than 32, the point *xc* is chosen from [33-64] randomly and $f(xc)$ is second minimum value and $f(xc)$ is larger than $f(xb)$. When $f(xa)$ is global minimum value, $f(xc)$ is second maximum value at [33-64] and $f(xc)$ is smaller than $f(xa)$. The point *xd* is also chosen from [65-128] which is shown in Fig.4(a) and Fig.4(b). In this way, 64 points are chosen from 128 points randomly after dividing into 32 parts. Then we can obtain the 64 points which take charge of maximum or minimum values.



(a) 2nd dividing

(b) 3rd dividing

Fig. 4. How to choose max. and min. points

(ii) The global maximum and minimum value of creating curve is restricted to 100 and 0. The average of difference between the maximum and minimum values for each part, *average*(n-1) which is divided n-1 times can be obtained from Hurst exponent. When $f(xa)$ is maximum value (=100) and $f(xb)$ is minimum value (=0). Then $f(xc)$ and $f(xd)$ are restricted as follows:

$$
\begin{aligned}
&f(xa)-f(xb)=average(0)=100\\
&\{f(xa)-f(xc)+f(xd)-f(xb)\}/2=average(1)\\
&f(xd)-f(xc)=2average(1)-average(0)
\end{aligned}
\tag{2}
$$

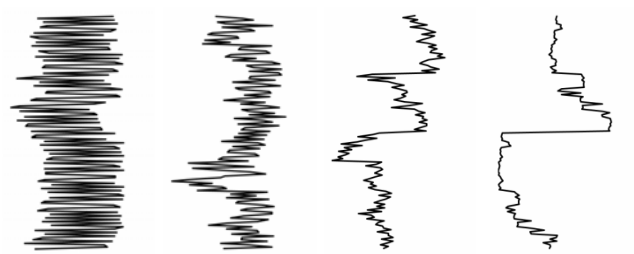
Here, $f(xa)>f(xd)>f(xc)>f(xb)$

Secondly, $f(xe)$ to $f(xh)$ are restricted as follows:

$$\begin{aligned}&\{ f(xa)-f(xe)+f(xf)-f(xc)+f(xd)-f(xg)+f(xh)-f(xb)\}/4\\&=average(2)-f(xe)+f(xf)-f(xg)+f(xh)\\&=4average(2)-average(0)-2average(1)+average(0)\\&=4average(2)-2average(1)\end{aligned} \tag{3}$$

Here, $f(xa)>f(xf)>f(xe)>f(xc)$ and $f(xd)>f(xh)>f(xg)>f(xb)$

Then, maximum and minimum values in each part are decided by the same way as above. Some generated curves by using the proposed system are shown in Fig.5. And, Fig.6 shows the Hurst exponents of generated curve based on the specific it. From this figure, the proposed method can supply curves based on the Hurst exponents.



He =0.1 0.3 0.6 0.9

Fig. 5. Examples of generated curves

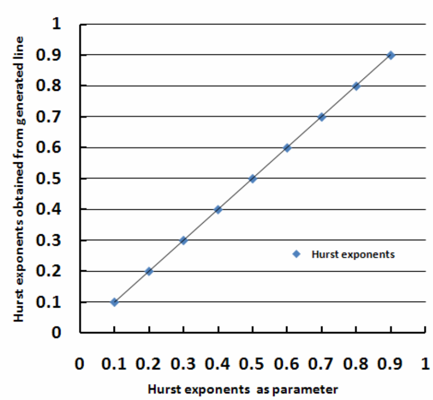


Fig. 6. Relation between Hurst exponents of curves and parameters

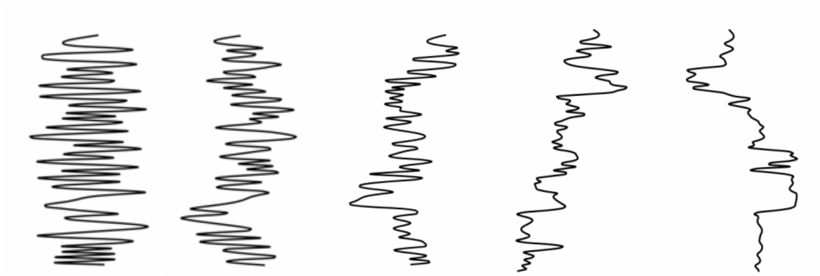
3 Human Feeling Complexity

3.1 Experimental Object

It is difficult to determine the human feeling complexity for a curve absolutely. Therefore, the sensory evaluation by paired comparison using some generated curves made by the proposed method, and relative complexities are resolved by statistical analysis.

3.2 Experimental Method

The curves as shown in Fig.7 were used in the sensory evaluation by paired comparison. Two out of five curves were presented to a subject at the same time, and the subject watched the curves. Then the subject was urged to answer which curve he or she felt more complex. The subjects are 20 students in their twenties.



$A_{0.172}$ $B_{0.312}$ $C_{0.452}$ $D_{0.591}$ $E_{0.741}$

Fig. 7. Generated curves used in sensory evaluation

3.3 Experimental Method

Table 2 shows the results of paired comparison. The number in each table shows that of subjects who felt the line curve was more complex than the row curve.

Table 2. Result of paired comparison

	$A_{0.172}$	$B_{0.312}$	$C_{0.452}$	$D_{0.591}$	$E_{0.741}$	Total
$A_{0.172}$		12	13	11	11	47
$B_{0.312}$	8		15	14	18	55
$C_{0.452}$	7	5		12	18	42
$D_{0.591}$	9	6	8		16	39
$E_{0.741}$	9	2	2	4		17

The Bradley-Terry model is assumed to evaluate the complexities of curves quantitatively, defined as follows [2];

$$P_{ij} = \frac{\pi_i}{\pi_i + \pi_j} \tag{4}$$

$$\sum_i \pi_i = const.(=100) \tag{5}$$

Where π_i : intensity of i,
P_{ij} : probability of judgment that i is more complexity than j.

π_i shows the intensity of complexity of the curve i. The model enables to determine the complexity based on the paired comparison. Here, the maximum likelihood method is used to solve the π We obtain $\hat{\pi}_i$ by the following formula where $\hat{\pi}_i^0 (i = 1,2,3)$ were used as initial values,

$$\hat{\pi}_i = \frac{T_i}{\sum_{j(\neq i)} \frac{N}{\hat{\pi}_i^0 + \hat{\pi}_j^0}} \tag{6}$$

N is the number of objects, and T_i is the total number of i's win. Then $\hat{\pi}_i$ is scaled up or down to satisfy the next formula.

$$\sum_i \hat{\pi}_i = K \tag{7}$$

where K is 100.

$$\hat{\pi}_i^1 = \frac{K\hat{\pi}_i}{\sum_i \hat{\pi}_i} \tag{8}$$

We iterated the series of calculation until $\hat{\pi}_i$ was settled. The result of B-T models is shown in Fig. 8.

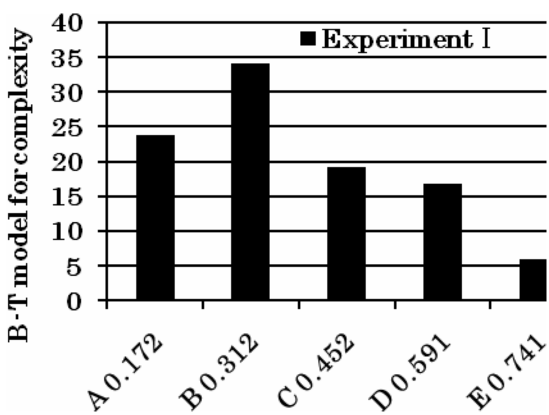
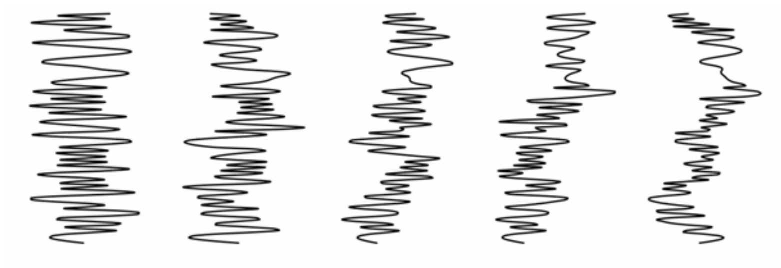


Fig. 8. Bradley-Terry model for Experience I

However, the curve A is more complex than B from the viewpoint of mathematics, the subjects tend to feel that B is more complex than A. This indicates that there is possibility of disagreement between the human feeling complexity and the mathematical complexity.

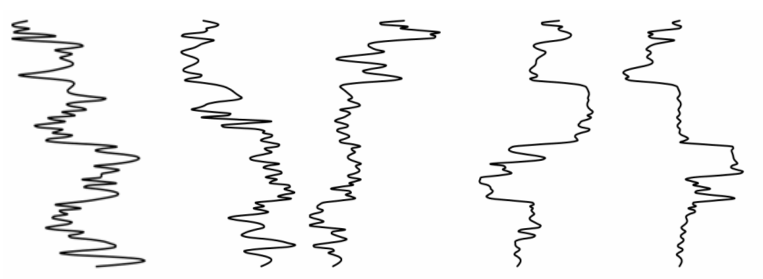
3.4 Experimental II

As a result of Experiment I, there is possibility that there is the disagreement between the human feeling and mathematical complexity when the Hurst exponents are small comparatively. Therefore, two experiments were performed in which Experiment II-1 used the curves which have small (<0.5) Hurst exponents, and Experiment II-2 used the curves which have large (>0.5) Hurst exponents comparatively. Fig.9 and 10 show the curves and their Hurst exponents used in Experiment II-1 and II-2.



$F_{0.298}$ $G_{0.363}$ $H_{0.407}$ $I_{0.423}$ $J_{0.459}$

Fig. 9. Generated curves used in Experiment II-1



$K_{0.527}$ $L_{0.579}$ $M_{0.634}$ $N_{0.683}$ $O_{0.744}$

Fig. 10. Generated curves used in Experiment II-2

The sensory evaluation by paired comparison was performed as same as Exp.I. Table 3 shows the results of paired comparison in Experiment II-1 and II-2. The Bradley-Terry models are also assumed to evaluate the complexities of curves quantitatively, and these results are shown in Fig. 11.

As same as the result of Experiment I, there are some disagreements between the human feeling and mathematical complexity in both Experiment II-1 and II-2. However the difference between them in Experiment II-2 is slight, it between them in Experiment II-1 is not so small.

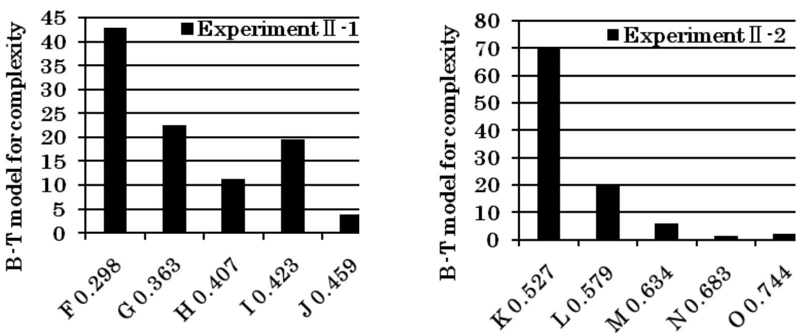
Table 3. Result of sensory evaluation in Experiment II

(a) paired comparison in Experiment II-1

	$F_{0.298}$	$G_{0.363}$	$H_{0.407}$	$I_{0.423}$	$J_{0.459}$	Total
$F_{0.298}$		17	15	13	16	61
$G_{0.363}$	3		15	13	17	48
$H_{0.407}$	5	5		6	17	33
$I_{0.423}$	7	7	14		17	45
$J_{0.459}$	4	3	3	3		13

(b) paired comparison in Experiment II-2

	$K_{0.527}$	$L_{0.579}$	$M_{0.634}$	$N_{0.683}$	$O_{0.744}$	Total
$K_{0.527}$		17	18	19	19	73
$L_{0.579}$	3		18	19	17	57
$M_{0.634}$	2	2		18	15	37
$N_{0.683}$	1	1	2		8	12
$O_{0.744}$	1	3	5	12		21

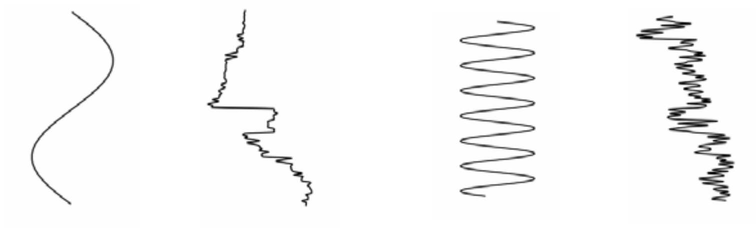


(a) Bradley-Terry model for Exp. II-1 (b) Bradley-Terry model for Exp.II-2

Fig. 11. Bradley-Terry models for Experiment II

4 Discussions

Some subjects pointed out the curves which they felt the periodic round were not felt so complex. Fig. 12 shows some generated curves which have same Hurst exponents as the sin curve. It is obvious that generated curves are felt more complex than sin curves. It is indicated that the curve which has periodic round is not so complex as its Hurst exponents by human feeling.



(a) sin curve (He=.107) (b) generated curve (He=.107) (c) sin curve (He=.480) (d) generated curve (He=.480)

Fig. 12. Sin curve and generated curve which have same Hurst exponent

5 Conclusions

A complexity of curve is defined by mathematical method absolutely. Hurst exponent in fractal geometry can indicate the complexity of a curve as a number which is between 0 and 1. On the other hand, the complexity of curve by human feeling cannot be defined absolutely, but relatively. In this paper, we have tried to compare the human feeling complexity to mathematical complexity by sensory evaluation and statistical analysis.

As a result of experiment, there is possibility that there is the disagreement between the human feeling and mathematical complexity when the Hurst exponents are small comparatively. Moreover, the periodicity of curve may effects on the human feeling complexity.

References

1. Okubo, M., Watanabe, T.: Sensory Evaluation of the Preference of 3D Shape in Virtual and Real Environments. In: IEEE International Workshop on Robot and Human Communication (RO-MAN 1997), pp. 460–464 (1997)
2. Hori: Brain and Emotion. Kyoritsu Press (1991) (in Japanese)

How Do Real or Virtual Agent's Body and Instructions Contribute to Task Achievement?

Yugo Takeuchi and Hisashi Naito

Shizuoka University,
3-5-1 Johoku, Naka-ku, Hamamatsu, Shizuoka 4328011 Japan
takeuchi@inf.shizuoka.ac.jp

Abstract. In face-to-face communication, we can confidently communicate through our bodies. Recently, agents have widely surfaced as existences that interact with humans, and several studies have investigated the formation of social relations with such agents. This research focuses on 'sharing an environment' when humans communicate with humans, and we adapt it to human and agent interactions. We examined the effects of sharing an environment in two cooperative task experiments in real/virtual worlds and found that reception frequency for information from the agent with a body in the real world is significantly higher than the reception frequency for information with the agent with a body in the virtual world. In conclusion, our results suggest that sharing the same environment through a body affected the reliability of the information.

1 Introduction

Since CG characters and robots have become communication objects for people, many studies have addressed how humans should associate with these existences that previously didn't exist. However, many of the objects that were treated by those researches have not infiltrated our lives yet. On the other hand, things already exist whose behavior reflects the user's purpose through interaction with that person. Such driving navigation machinery as car navigation systems, which are known for advanced information and communication machinery, masks the possibility of the realization of various future interactions.

In guidance by car navigation systems, such concrete numerical values as "turn left in 100 meters" are used. This is an overlook viewpoint that uses concrete numerical values. On the other hand, in guidance by humans, landmarks such as "turn left just before the Korean restaurant" are more common, and such deictic phrases as "turn right at that crossing" are used. Human instructions are presumably easier to understand because the driver and the guide share viewpoints. In addition, the navigator is sitting next to the driver, and they share a place where both of their bodies are located. Accordingly, a sense of security is created in the driver. The result that accuracy of action and understanding degree for information presentation of the information receiver improves is reported by sharing of information and sharing space through a body in human-human communication [1].

A variety of personification agents and robots are beginning to permeate society due to improved computer capacity and advances in robot technology. In the future,

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 142–151, 2011.

as agents become more prevalent, situations are expected to increase for which we can use agents; new kinds of communication will be realized by cooperation with them. From such a background, in this study, we concentrate on a method of presenting information and the body's role in information processing activities based on the real world and discuss the directionality of new interaction designs between humans and agents.

2 Real-World Oriented Interaction

Real-world oriented interactions communicate between humans and computers with a real-world base. Most of our various activities are done in the real world. Reality is improved by being connected to computers in the real world and relating computer information to the real world; more natural interaction can be designed.

The information that can be offered from such a system to human beings exists in a database and in virtual space and is fundamentally isolated from the real world. When information is presented to a user, who thereby becomes the information receiver, the information presented is from an environment different from the real world in which the user exists; the information is isolated from the environment (Fig. 1).

However, with the continued development of information technology, the relationship between the environment and such information has changed. With the development of network technology, miniaturization, and improvements in the appearance of such new input methods as voice recognition, motion sensors, and touch panel displays, information exists in the same real world in which the user exists. Thus we can predict that real-world oriented interaction can be realized through the presentation of information from identical environments (Fig. 2).

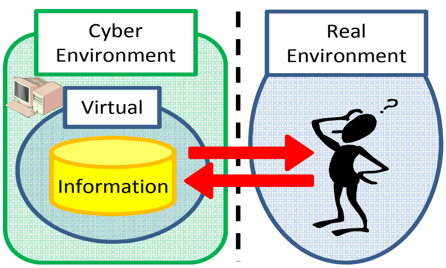


Fig. 1. Image of conventional presentation of information

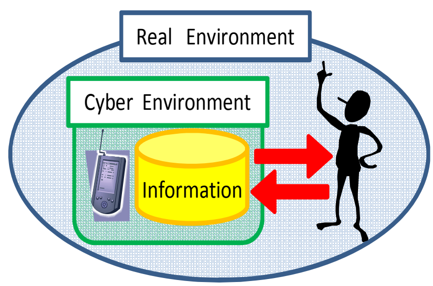


Fig. 2. Image of real-world oriented interaction

For example, consider the following two examples. The first is the technology of tangible user interfaces with which we can directly touch information that has no form [2]. The second is such technology as Augmented Reality and Mixed Reality that put the virtual object and the information generated by a computer in the real world. Treating information as part of our surrounding environment allows information to be exchanged without being conscious of a computer. This point is an example of real-world oriented interaction.

First, we discuss sharing of information and sharing of space through bodies as examples of real-world oriented interaction between humans. Next we discuss the agent reliability that is treated in this study.

Sharing information. In communication with other people, a person reduces his speech by using various kinds of pre-linguistic information such as demonstrative pronouns. We can carry out this method of reduction for objects to which we can point and handle directly by actually manipulating or touching them. In addition, we can carry out this reduction for objects with which we indirectly deal by just seeing or having them close. When we use our bodies to conduct such operations, reduction is crucial that uses indirect ways of dealing with objects, such as visual information.

Regarding such visual information, Nishikawa et al. suggested that smooth communication is possible when people share direct 'visual information' of the kind of visual information the person being talked to is receiving [4]. Based on this finding, we can say the following about the visual information used when speech is reduced. 'What the other person in a conversation is looking at' is 'something that expresses his visual information.' In other words, sharing information, such as the other person's visual information related to the 'present environment,' is an important point for smooth communication.

Sharing of space through bodies. In face-to-face communication such nonverbal information as gestures and nodding produces an entrainment of biological rhythms and makes communication more efficient [5]. For example, we draw a partner into conversation by nodding. In this way, communication using body language is human-human communication. Embodiment shares time and space through each other's bodies and supposes that people share knowledge and situations with a communication partner. Trust in one's partner also emerges. We propose a communication system that incorporates these features [6].

3 Cooperative Task Experiment in Real World: Experiment 1

3.1 Purpose

This experiment examines hypotheses about the differences of agent bodies and whether information offered by agents corresponds to the environment influence of human evaluations of agents. We set a treasure hunt game as a cooperative task problem in the real world. First, we observed whether the information of the real-world agent or the virtual-world agent was more trusted. Second, we observed which agent is more trusted: the agent providing information that corresponds to the environment or the one providing information whose correspondence with the environment is vague.

3.2 Participants and Task

45 university students from 18 to 25 years old participated. Subjects performed a simple treasure hunt game with an agent (Fig. 3) and began from the starting point shown in Fig. 1. The agent presents route selection information to arrive at the diverging point. The content of the agent instructions about the route selection to the diverging

point is assumed to be correct. When subjects go through the maze and arrive at the correct place (one among T1, T2, T3, and T4 in Fig. 3), they find three paper cups where the chest is assumed to be located. The treasure is defined by a constant sound that originated from one of the three cups. Subjects used a parabolic reflector to understand to some extent which paper cup is producing the sound. The agent gives information about which of the three is the treasure. A number is written on the paper cup, and the agent provides information by saying a number. We compared cases where the agent showed the correct number with the incorrect number to observe the behavior differences of subjects when specifying treasure from the three treasure chests. Fig. 4 shows the treasure specified with a parabolic reflector.

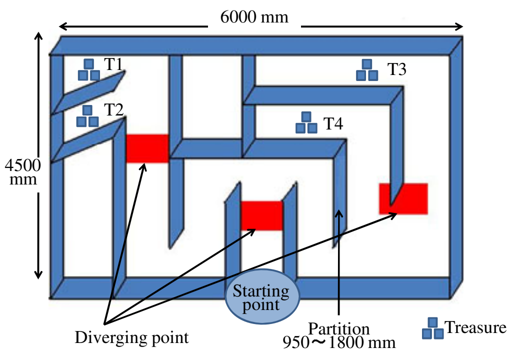


Fig. 3. Overview of task field

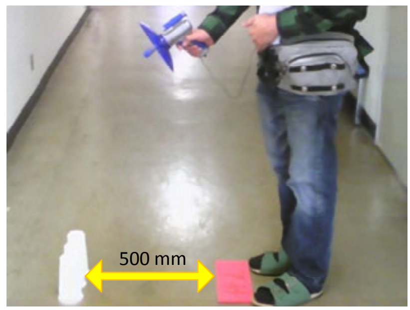


Fig. 4. Subject identifies which sound is emitted from among three paper cups

3.3 Conditions

There are two factors: agent's body (body factor) and the relation between the information and environment (environment factor) (Table 1). There are three body factors: robot (Fig. 5), CG (Fig. 6), and the computer. There are two environment factors: corresponding and vague. The corresponding information is the instructions about the concrete route selection (example: That is the right route.). The vague information is not instructions about the concrete route selection but rather the position instructions of the treasure with a direction and a distance (example: It is 100 cm in the 12 o'clock direction.). There are 3 x 2 = 6 conditions.

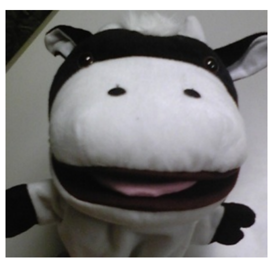


Fig. 5. Robot agent

Fig. 6. CG agent

The body factor is between the subjects, and the environmental factor is within the subjects. Subjects played the treasure hunt game four times, twice in each environmental factor. First, we lowered the degree of difficulty of choosing from three by increasing the volume, and agents provided correct treasure information to form a

Table 1. Conditions in Experiment 1

		Environment factors	
		Corresponding	Vague
Body factors	Robot	Condition A	Condition B
	CG	Condition C	Condition D
	Computer	Condition E	Condition F

relationship of mutual trust with the participants. In the second, we raised the degree of difficulty by reducing the volume, and agents gave incorrect treasure information. Here we observed whether subjects believed the sound or themselves.

3.4 Hypothesis

This experiment was performed based on a hypothesis that people will willingly trust information from objects that share an environment (Hypothesis 1) and information that corresponds to the environment (Hypothesis 2). This experiment task played a treasure hunt with agents who presented correct information about the route selection and information about the treasure's location. When an agent gives incorrect information to identify the treasure, we expect that subjects will trust hints from the robot agent more and obey the instructions, and the subject will trust hints more from agents that provided corresponding information when the subject is choosing a route. According to this hypothesis, we predict the following:

- In the body factor, the reception frequency for information from the robot agent is significantly higher than the reception frequency of the two other levels (Conditions A, B > Conditions C, D, E, F).
- In the environment factor, the reception frequency for the corresponding information is significantly higher than the reception frequency for the vague information (Conditions A, C, E > Conditions B, D, E).

3.5 Results

We assigned 15 people in each body factor. The results for the number of times a subject obeyed when the agent gave incorrect information appear in Table 2 and Fig. 7. In the body factor, subjects readily obeyed the instructions of the robot agent more than the instructions of the CG agent and the computer, and there is hardly a difference between the CG and the computer. In the environment factor, subjects easily obeyed the corresponding information for the robot condition. But there was no meaningful difference.

Next we regarded the case where the subject obeys the instruction as 1 point and doesn't obey as 0 point and performed a two-way repeated measure analysis of variance. Analysis revealed that the simple main effect of the body factor is significant ($F_{(2,42)}$=3.57, p<.05). A simple main effect of the environment factor was not seen ($F_{(2,42)}$=0.79, n.s.). In addition, we performed a multiple comparison with the LSD method, and the averages of the robot condition were significantly larger than each average of the other two conditions (MSe=0.2523, p<.05).

Table 2. Number of times instructions were obeyed in Experiment 1

	Robot		CG		Computer		Total
	Corresponding	Vague	Corresponding	Vague	Corresponding	Vague	
Obey	13	9	7	6	6	7	48
Disobey	2	6	8	9	9	8	42
Total	15	15	15	15	15	15	90

3.6 Consideration

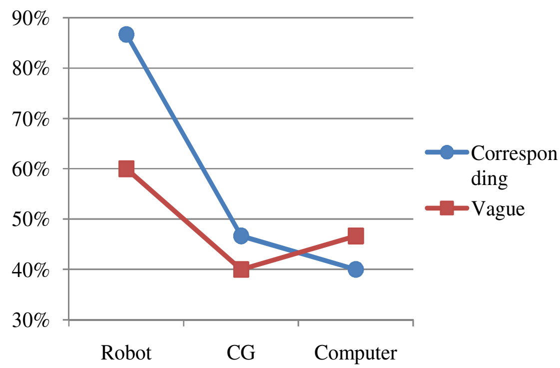


Fig. 7. Percentage who followed instructions in Experiment 1

Our experiment results show that the existence of a body in the real world improves the agent evaluation; Hypothesis 1 was supported. This result resembles the preceding study described in Section 2.2. In the real-world task, humans easily trusted agents who exist in the real world. Because the subject actually had a robot, the vibration of the robot's movement may have strongly impressed subjects with the agent's existence.

On the other hand, no significant difference was seen for the environment factor; Hypothesis 2 was not supported. The following is offered as the reason. First, since we experimented on two conditions of the environmental factor with one subject with the same agent, the agent may not have been distinguished between the two conditions. Second, vague information produced a hesitation of the route choice and an assumption that the agent reliability had deteriorated. However, a correct route choice was enabled by vague information because we repeatedly experimented with a comparatively easy maze; no difference appeared.

In addition, because a difference was only seen as a tendency between the environmental factors for the robot, agent reliability will increase most by exchanging the information that corresponds to the environment, assuming that the body exists in the real world.

4 Cooperative Task Experiment in Virtual World: Experiment 2

In Experiment 1, all humans, agents, and tasks were unified in the real world. If the task were something in the virtual world, we believe that the place where agents exist is important because humans exist in the real world, but the task exists in a virtual world. Therefore we experimentally inspected the agent's physicality in a virtual world in a treasure hunt game that resembled Experiment 1.

4.1 Purpose

This experiment examines the hypotheses about the differences of agent bodies and the methods of presenting information that influence human evaluations of agents. We set a treasure hunt game as a cooperative task problem in the virtual world. First, we observed whether the information of the real-world agent or the virtual-world agent is more trusted. Second, we observed whether the absolute information that uses a direction was expressed in the virtual world for the screen direction or the relative information that uses a direction against the screen direction.

4.2 Participants and Task

34 university students from 18 to 25 years old participated and performed a simple treasure hunt game with agents using a maze in an immersive, virtual three-dimensional space. Agents led subjects to the treasure's location. Subjects arrived at the diverging point when they went through the maze. At the diverging point, first we showed correct information about in which direction the treasure exists against the present position by sound. The sound information, which is the direction where the treasure exists, may not show the correct route. Second, agents presented route selection information. The content of the agent instructions about the route selection to the diverging point are assumed to be correct. We observed which information the subject believed when the sound information did not agree with the information by the agent. Fig. 8 shows a experiment system screen that used a virtual agent.

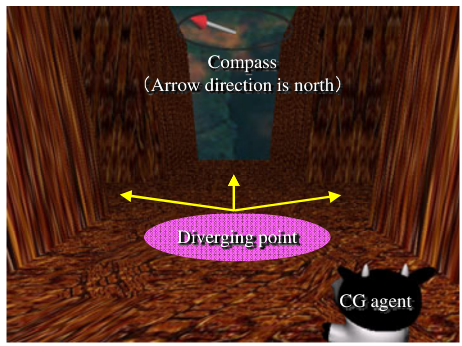


Fig. 8. Experimental task

4.3 Conditions

There are two factors: agent's body (body factor) and relation between information and environment (environment factor) (Table 3).

There are two body factors: robot and CG. There are two environment factors: relative and absolute. The relative information is the instructions that included a direction against the direction of the subject eyes (screen direction) (example: That is the 9 o'clock direction.). Absolute information is the instructions that provided a direction expressed in the virtual world (for example: That is northeast). There are 2 x 2 = 4 conditions.

The relative instruction is the information presentation method based on the subject viewpoints (environment) because the instructions are based on information provided from the subject viewpoints. Therefore, we assume that the cognitive load is light. Meanwhile the absolute instruction method does not correspond directly with the environment of the subjects because it is based on the virtual world's environment. Therefore, we assume that the cognitive load grows because the instruction content must be associated with the virtual world's environment.

The body factor is between subjects, and the environmental factor is within subjects. Subjects performed the treasure hunt game four times, twice in each environmental factor. In the first game, we led subjects accurately to the place with the treasure by matching the sound information to the agent instructions. Agents provided correct treasure information to form a relationship of mutual trust with the subjects. In the second game, we changed the sound information so that it differed from the agent instructions. Here, we observed whether subjects believed the sound or the agent.

Table. 3. List of conditions in Experiment 2

		Environment factors	
		Relative	Absolute
Body factors	Robot	Condition A	Condition B
	CG	Condition C	Condition D

4.4 Hypothesis

This experiment was performed based on a hypothesis that people will easily trust information from objects that exist in the same environment as the task (Hypothesis 1) and information from objects that present information based on one's own environment (Hypothesis 2). When instructions by agents and the sound information are different, we expect that subjects will trust the hints from the GG agent more and obey the instructions and subjects will trust hints more from the agent that provided the relative information. According to this hypothesis, we predict the following:

- In the body factor, the reception frequency for the information from the CG agent is significantly higher than the reception frequency for the information from the robot agent (Conditions C, D > Conditions A, B).
- In the environment factor, the reception frequency for the relative information is significantly higher than the reception frequency for the absolute information (Conditions A, C > Conditions B, D).

4.5 Results

We assigned 17 people in each body factor. The results for the number of times subjects obeyed when the instructions by agents and sound information were different appear in Table 4 and Fig. 9. In the body factor, subjects readily obeyed the instructions of the robot agent more than the instructions of the CG agent. In the environment factor, subjects easily obeyed the relative information more than the absolute information.

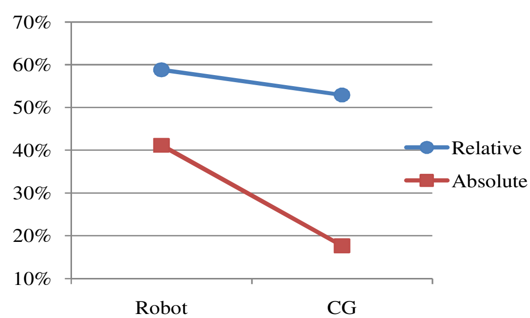


Fig. 9. Percentage who followed instructions in Experiment 2

Next we performed a two-way repeated measure analysis of variance using the same method as Experiment 1. No significant difference was observed in the body factor ($F_{(1,32)}$=1.52 ,n.s.). The simple main effect of the environment factor is significant ($F_{(1,32)}$=5.31, p<.05).

Table 4. Number of times instructions were obeyed in Experiment 2

	Robot		CG		Total
	Relative	Absolute	Relative	Absolute	
Obey	10	7	9	3	29
Disobey	7	10	8	14	39
Total	17	17	17	17	68

4.6 Consideration

Our experiment results show that the place where agents exist does not influence the reception frequency of the subject for the agent instructions; Hypothesis 1 was not supported. Because a tendency was observed of easily obeying the instructions of the robot agent, we assumed that the impression between the robot and the CG was different, even though the difference of agent embodiment did not directly influence subject behavior. We believe that the quality of the information for the receiver of the instructions changes even if the instruction content remains the same, regardless of the task environment, because the body of the agent exists in the same environment as the human.

On the other hand, the results show that the reception frequency for the relative information is greater than for the absolute information; Hypothesis 2 was supported. The method that uses absolute coordinate based on direction in the virtual world is accompanied with a high cognitive load, so it is necessary to check which direction we are facing. On the other hand, interpretation was quick of the method based on the direction the screen faces. Subjects apparently avoided interpreting the absolute instructions with cognitive load and followed the relative instructions that they could interpret intuitively.

5 General Discussion and Conclusion

This study examined whether the place where agents exist and the relation between information and environment affected the reliability of information presented by agents through two real and virtual world tasks. Through this study we showed the following:

- In the real-world task, the reception frequency for information from the agent increased since the bodies of both humans and agents exist in the same environment: in the real world.
- In the virtual world task, the reception frequency for information from agents tended to increase since the bodies of both humans and agents exist in the same environment: in the real world.

- The reception frequency for information from agents increased because the instructions corresponded to the environment of the human's viewpoint.

In the agent's embodiment, we revealed the utility of the substance agent in the real world task from Experiment 1. This utility was also suggested in the virtual world task from Experiment 2. These suggest that when we work cooperatively with agents, we feel as if we are sharing the same environment with them because their bodies exist in the real world where humans exist regardless of the task environment. Even if the information context is identical, perhaps only a different form of agent bodies changes the quality of information to the receiver of information. In these two experiments, humans existed in the real world. The utility of virtual agents is expected when humans exist in the virtual world.

In the relation between information and environment, we revealed that the method of presenting information that corresponds to the human environment, especially the view environment, raises the reception frequency for information. This suggests that human cognitive load may be reduced by concealing information in environments when humans acquire it through interaction with the outside world.

References

1. Kawamura, S., Yamada, Y., Itoh, T., Takeuchi, Y.: An Empirical Study of Identificating Reliabile Messages in Real World Interactions. In: Proceedings of 4th Biennial Workshop on DSP for In-Vehicle Systems and Safety, DSP35 (2009)
2. Bonanni, L., Ishi, H.: Stop-motion prototyping for tangible interfaces. In: Proceedings of the 3rd International Conference on Tangible and Embedded Interaction, TEI 2009, Cambridge, United Kingdom, February 16-18, pp. 315–316. ACM, New York (2009)
3. Nishikawa, a., Kijima, S., Katoh, T., Maruya, M., Yamaguchi, A., Miyazaki, F.: Mutual View Sharing System for Real-Time Tele-Communication Systems and Computers in Japan, vol. 37(14), pp.73–86 (2006)
4. Watanabe, T.: E-COSMIC: Embodied Communication System for Mind Connection. In: Proc. of the 13th IEEE International Workshop on Robot-Human Interactive Communication (RO-MAN 2004), pp. 1–6 (2004)
5. Ishibiki, C., Nakajima, Y., Matsumoto, D., Miwa, Y.: Expression of Existence Supporting Sharing Interspatial Distance-"Maai": Development of Roving Object Integrating Partner's Shadow and Video Image. Journal of Robotics and Mechatronics 17(3), 310–317 (2005)

Interaction Mediate Agent Based on User Interruptibility Estimation

Takahiro Tanaka and Kinya Fujita

Graduate School, Tokyo University of Agriculture and Technology
2-24-16 Nakacho, Koganei, Tokyo 184-8588, Japan
{takat,kfujita}@cc.tuat.ac.jp

Abstract. In recent years, the scope of users receiving information from information systems has been increasing. However, the timing of interruption is not controlled in most of these systems. In this paper, we propose a secretary agent to mediate interactions between users and others. The agent estimates user interruptibility from PC operation records and head motions. Moreover, the agent requests interaction from others by controlling the avatar's gaze, joint attention, and mutual gaze on the basis of the interruptibility. We confirmed the usefulness of the mediation based on the estimation of user interruptibility.

Keywords: interruptibility, estimation, human–agent interaction, application switching.

1 Introduction

In recent years, the scope of users receiving information from information systems has been increasing with the popularization of the Internet and ubiquitous computing environments. Although the frequency of interruptions from systems such as mail alerts, instant messaging tools, and software agents has been increasing, the timing of such interruptions is not controlled in most information systems. Previous studies [1] have indicated that frequent interruption without considering the user's status fragments the user's work time and decreases productivity. Therefore, automatic status estimation and ambient interruption methods are expected to assist in avoiding unintended interruptions.

One potential method for estimating a user's online status is by monitoring PC operations or by using sensors [2-6]. However, these physical activity indices do not capture all instances of the intellectual activities that should not be interrupted, because intellectual activities do not always have observable outputs. Another approach is to estimate the breakpoint of a task [7, 8]. At the breakpoint, the user's interruptibility is expected to increase instantaneously. However, this approach requires task-structure analysis to determine the interruptibility level of the breakpoint.

We considered focused application switching (AS) as a breakpoint in PC work. Our experimental results demonstrated that the interruptions at AS are significantly more acceptable than those during continuous work. Finally, we proposed a user interruptibility estimation method for AS timing [12].

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 152–160, 2011.

In this paper, we propose a secretary agent to mediate interactions between the user and other users or systems, in order to start accepting interactions with others, based on the interruptibility estimation method. The agent gently requests interactions from the user by using gazing actions (joint attention and mutual gaze) and leads the user to confirm the request on his/her own. We experimentally confirmed the usefulness of the agent for estimation and mediation between users and others.

2 Related Studies

2.1 Estimation of User Business/Interruptibility

There are several related studies regarding user-status estimation by various techniques; e.g., counting the keystrokes or mouse clicks and using various sensors placed in the work space or on the users [2-6]. These methods are expected to adequately estimate the user's status when the status has observable physical activity indices. However, intellectual activity such as deep thinking has no observable output. Moreover, it is practically difficult to place various sensors on the user or in the work space.

On the other hand, some studies have reported a relationship between interruptibility and breakpoints [7, 8]. At a breakpoint, the user's interruptibility is expected to increase instantaneously even if the work requires intellectual activity. Furthermore, it has been suggested that the kind of the breakpoint affects interruptibility. However, this method requires an analysis of the task structure to determine the level of the breakpoint. Multi-tasking during PC usage is also a potential difficulty in the application of the task-structure based method. Therefore, it is still a challenge to distinguish the level of interruptibility at a breakpoint.

Moreover, the interruptibility expression must be intuitive in order to reduce the user's cognitive load. Some studies have examined ambient awareness using nonverbal information. Movements of the head, face, and body were determined to be effective in expressing nonverbal behavior [9]. The expression of interruptibility using nonverbal information through an avatar appears promising for the intuitive recognition of a user's status.

2.2 Study of Application Switching for User Interruptibility Estimation

We used focused application switching (AS)—transition of the active application window—as a breakpoint in PC work [12]. AS is considered to be the user's intentional switching of the working space or working target. Therefore, the user's concentration at AS is expected to be instantaneously weakened compared to that during continuous work. Moreover, AS commonly and frequently occurs in PC work and is easily detected. Thus, AS offers scope for information presentation with less risk of task disturbance.

To examine this assumption, we experimentally collected and analyzed PC operation records and subjective interruption evaluation logs. Experimental results demonstrate that interruptions at AS are significantly more acceptable for users than those during continuous work ($p < 0.01$, t-test).

Therefore, we analyzed the relationship between interruptibility scores and the indices calculated from the operation records, which were expected to reflect the interruptibility at AS. Finally, we defined the estimation rule on the basis of co-occurring indices at the same AS, and decided to judge the interruptibility using three levels in a way similar to the previous works, for practical utility. Figure 1 shows the experimental results of the interruptibility estimation using our proposed method. The results demonstrate that the proposed estimation method allows information systems to reduce the risk of serious work disturbance due to interruptions.

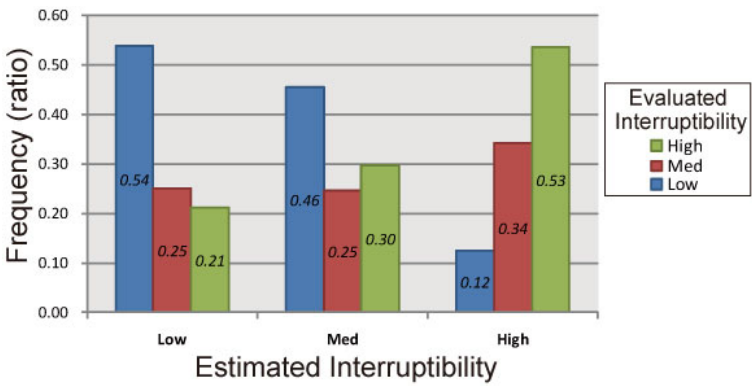


Fig. 1. Results of interruptibility estimation using three levels

3 Secretary Agent

We proposed an interaction initiation method with acceptable timing by using an agent to mediate the interaction based on the user-interruptibility estimation. In this section, we explain the concept of agent mediation and details of agent architecture.

3.1 Concept of Interaction Initiation by Agent's Mediation

The above analysis indicates that interruptions at AS are significantly more acceptable for users than those during continuous work. Based on this analysis, we proposed the interruptibility estimation method. This method provides a controlling interruption timing that is useful to initiate interactions without causing problems of disturbance from online users or other information systems. However, as the estimation has an error, there is a risk of unintentional disturbance. Therefore, we need to consider a method that is robust against estimation errors.

Interruptions by pop-up windows or sounds may seriously disturb the user's work. Therefore, an ambient and natural presentation of the interruption is necessary to enable the busy user to not respond until a break [9-11]. Moreover, it is difficult to predict a highly interruptible time. Therefore, assistance from the information system is needed for smoothly starting the interaction at a high-interruptibility period.

In this research, the secretary agent estimates user interruptibility and receives interaction requests from a user or system in bulk. After receiving the request, the agent starts to convey it at a high-interruptibility period by using an ambient presentation. The presentation is based on actions in human daily life, "wait and see how the interaction partner works." In particular, the agent gently presents the request to the user by gazing actions and leads the user to start the interaction. Therefore, the agent's mediation causes a shift of positions between the interrupter and the interrupted user.

Figure 2 shows an example of dialog initiation by the secretary agent. The dialog initiation has the following steps:

1. **Request dialog:** If user A wants to start a dialog with user B, A sends a dialog request to B's agent through his/her own agent via the network. Then A awaits a message from B while doing his/her own work.
2. **Start appeal:** The dialog request by A is sent to the secretary agent of B, and the agent starts to present the dialog request unobtrusively by using gaze controls, joint attention, and mutual gaze. The agent continues to appeal to B according to his/her interruptibility.
3. **Notice:** If B notices the appeal and decides to start a dialog with A, B sends an instant message using the system to A. On the other hand, B is allowed to not respond to the request until a break, because A is doing his/her own work and is not able to recognize whether B has not noticed the appeal or is ignoring the request.

3.2 Overview of the Secretary Agent

Figure 3 shows the architecture of the secretary agent. The developed agent consists of two main functional components, the user's interruptibility estimation component and the request appeal component.

The interruptibility estimation component monitors the user's PC operation activity and head position using a web camera and then estimates user interruptibility using three levels based on our proposed method [13]. The estimated interruptibility is used to control the intensity of an appeal using a CG avatar. The request appeal component expresses the existence of an interaction request from others by the joint attention and

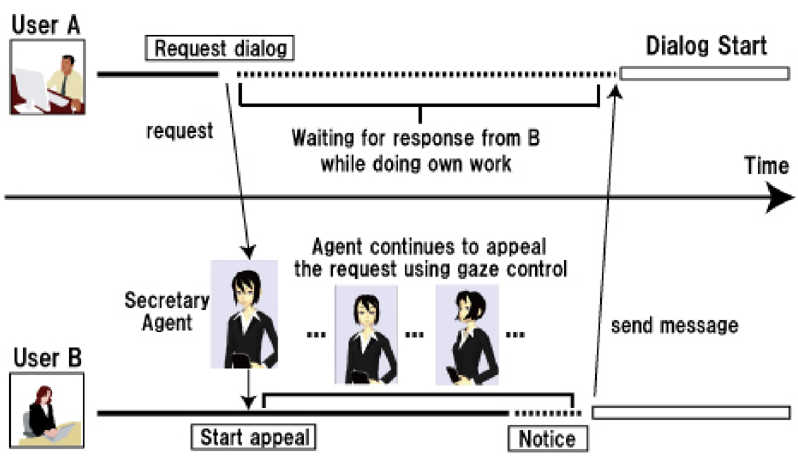


Fig. 2. Example of dialog initiation by the secretary agent's mediation

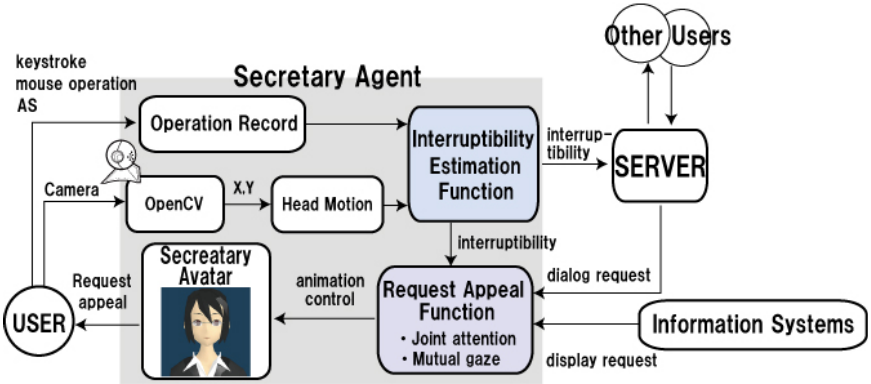


Fig. 3. Architecture of the secretary agent

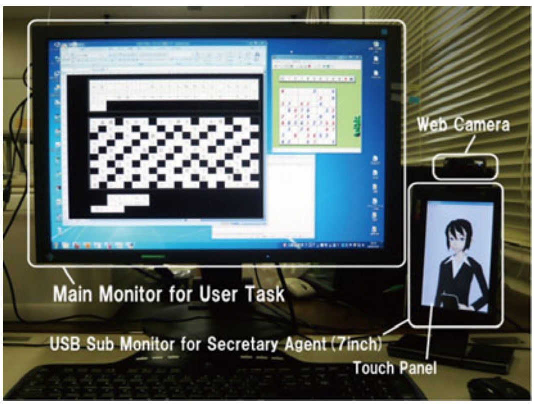


Fig. 4. Example of user environment

mutual gaze motions. The presentation strength and timing of the request are modified on the basis of the estimated interruptibility in order to attain pleasant and acceptable dialog initiation.

Figure 4 shows an example of user environment using the secretary agent. In this study, we propose using a USB sub-monitor to display the agent. Moreover, a web camera, which is used to capture the user's head position, is set up at the top of the sub-monitor. This not only improves the practicality, because it does not encroach upon a user's working area, but also reinforces the weak awareness of software agents. Moreover, a sense of intimacy and reliance on the agent are achieved easily because of the continuous interaction with the agent existing at the same location. We expect that a positive impression of the agent reduces the adverse effect due to the estimation effort.

Function I: User interruptibility estimation

The agent estimates the user interruptibility at AS timings in three levels based on our proposed method [13]. One difficulty in the previous research was that there were insufficient opportunities to interrupt at high interruptibility, only twice an hour. Therefore, in this study, we developed a method to estimate the interruptibility even during continuous work (not application switching: NAS).

To improve the accuracy of the estimation and to expand the estimation range of the target work, we considered using the user's head motion because it tends to reflect the user's motivation toward the current work [14]. To recognize the user's head motion, the agent captures the user's image using the web camera, and then detects the user's facial position and motion using OpenCV.

Function II: Ambient appeal of interaction requests

The agent presents a specific interaction request by using both joint attention and mutual gaze.

- **Joint Attention:** If the partner is in the NAS state, the agent occasionally observes the user's active window, which might be the user's current work space. The agent then appeals that the requested user is interested in the partner.
- **Mutual Gaze:** On the partner's AS, the agent gazes at the user who might be sitting in front of the monitor. At the time of AS, the interruptibility of the user will be reduced temporally, and so the agent uses a stronger presentation of the request, "face to partner," to influence the user to talk.

User interruptibility continually changes during the user's task. Therefore, the intensity of the appeal should be changed according to user interruptibility, or else the appeal would easily disturb the user's concentration even if it was ambient. Figure 5 shows a model of appeal intensity control for joint attention and mutual gaze based on an utterance attitude model [12]. The largeness and frequency of movement are

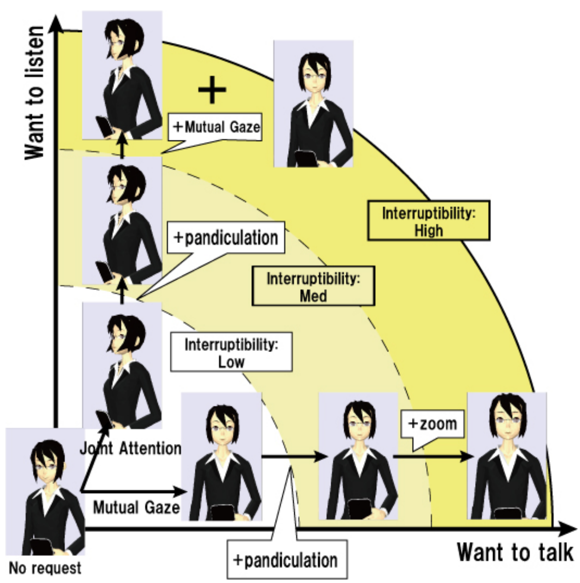


Fig. 5. Appeal intensity control for joint attention and mutual gaze

controlled on the basis of the user's interruptibility in three levels so as not to disturb the user's work. When the user's interruptibility is low, the agent appeals the request by using small movements and a low frequency. On the other hand, when the interruptibility is high, the agent frequently appeals the request by using large movements.

4 Experiment

The principal aim of this experiment is to confirm whether our interruptibility estimation method is useful for controlling presentation timing and whether the usage of a sub-monitor affects the impression of the secretary agent.

4.1 Procedure

In this experiment, the agent interrupts a subject performing given tasks in four experimental conditions. On being interrupted, the subject evaluates the interruptibility at the time. The evaluation is scaled from 1: "Absolutely uninterruptible" to 5: "No problem." Moreover, the subject evaluates his/her impression of the agent from three viewpoints by using a five-point scale: accuracy of interruption timing, tolerance of poorly timed interruptions, and motivation to use the agent. The subjects were eight university students who participated in this experiment in an environment shown in Figure 4. The subjects were assigned two tasks, a paper test and a puzzle. In addition, they were informed that the agent estimates the interruptibility and interrupts at highly interruptible times.

Two factors are crucial in this experiment. One is the control of presentation timing; the agent either interrupts the subject at random or estimates the interruptibility and interrupts at a highly interruptible time. The other is the position of the secretary agent; at the corner of the main monitor or sub-monitor. We set four experimental conditions as follows:

- A) Random timing and displaying in main monitor.
- B) Random timing and displaying in sub-monitor.
- C) Estimated timing and displaying in main monitor.
- D) Estimated timing and displaying in sub-monitor.

4.2 Results and Discussion

Figure 6 shows the evaluated interruptibility scores in random and estimation conditions. The average interruptibility of random interruptions was 2.5, whereas the average interruptibility of estimated interruptions was 3.6. Almost all random interruptions were during low interruptibility. On the other hand, the estimated interruptions were at high interruptibility. The interruptions at estimated timings were significantly higher than those at random. Thus, the results indicate the usefulness of the control of presentation timing based on the proposed method.

Figure 7 shows an interaction diagram of the four experimental conditions. The results of two-way factorial ANOVA without replication indicate that the level of interruptibility significantly affects the impression of the agent for all evaluations. Thus, it is considered that the interruption timing is very important for the impression of the agent because the agent has a well-defined purpose of controlling presentation timing.

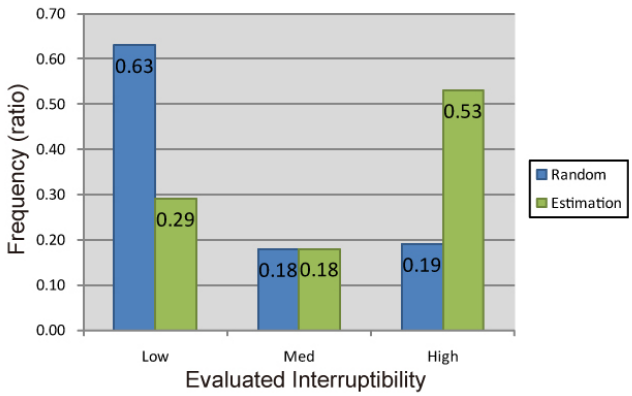


Fig. 6. Evaluated interruptibility: random vs. estimation

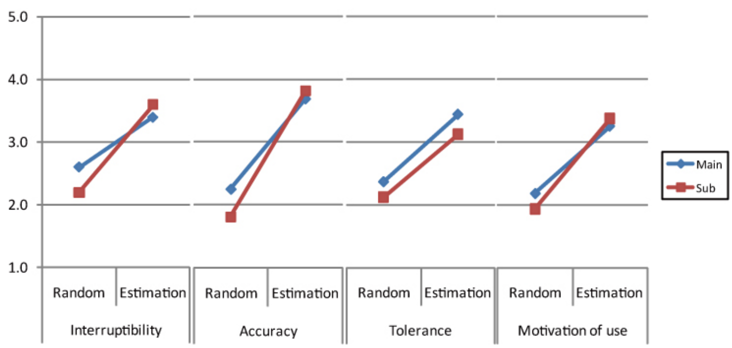


Fig. 7. Interaction diagram of evaluated scores

There was no significant difference regarding the display position of the agent. However, the subjects rated the agent displayed in the main monitor higher than that in the sub-monitor when evaluating the tolerance of a poorly timed interruption. At this point, we additionally interviewed the subjects regarding the reasoning behind this high rating. Some subjects assigned low scores for the tolerance because they expected improved agent performance with the use of particular devices and spaces such as a sub-monitor, and were disappointed because they did not observe any particular advantage of the device. In this experimental condition, we had turned off the request-appeal function because we attempted to evaluate the availability of the proposed estimation method. Therefore, the agent never moved during the experiment, even if it had a dedicated device and space. The user evaluates the agent according to not only its performance but also the cost or impression of the devices that comprise it.

5 Conclusions

In this paper, we proposed a secretary agent to mediate interactions between a user and other users or systems in order to start accepting interactions based on the interruptibility estimation method. We experimentally confirmed that our proposed estimation method is useful for controlling the agent's presentation timing. In the future, we shall experimentally evaluate the ambient appeal method by using gazing actions to interrupt without causing stress.

Acknowledgements. This research was supported by JSPS KAKENHI (21700130) and National Institute of Information Technology (NICT).

References

1. Bailey, B.P., Konstan, J.A., Carlis, J.V.: The Effect of Interruptions on Task Performance, Annoyance, and Anxiety in the User Interface. In: Proc of INTERACT 2001, pp. 593–601 (2001)
2. Honda, S., et al.: A Home Office Environment Based on the Concentration Degrees of Wokers: A Virtual Office SystemValentine. Trans. Info. Processing Society of Japan 39(5), 1472–1483 (1998) (in Japanese)
3. Minakuchi, M., Takeuchi, T., Kuramoto, I., Shibuya, Y., Tsujino, T.: An Automatic Estimation Method for Busyness at Deskwork. Trans. Human Interface Society 6(1), 69–74 (2004) (in Japanese)
4. Fogarty, J., Christensen, J.L.J.: Presence versus availability: the design and evaluation of a context-aware communication client. Int. J. Human-Computer Studies 61(3), 299–317 (2004)
5. Chen, D., Hart, J., Vertegaal, R.: Towards a physiological model of user interruptability. In: Baranauskas, C., Abascal, J., Barbosa, S.D.J. (eds.) INTERACT 2007. LNCS, vol. 4663, pp. 439–451. Springer, Heidelberg (2007)
6. Hudson, S.E., Fogarty, J., Atkeson, C.G., Avrahami, D., Forlizzi, J., Kiesler, S., Lee, J.C., Yang, J.: Predicting Human Interruptibility with Sensors: A Wizard of Oz Feasibility Study. In: Proc. SIGCHI Conf. on Human Factors in Computing Systems, pp. 257–264 (2003)
7. Czerwinski, M., Cutrell, E., Hirvutz, E.: Instant Messaging and Interruption: Influence of Task Type on Performance. In: Proc. of OZCHI 2000, pp. 356–361 (2000)
8. Iqbal, S.T., Bailey, B.P.: Leveraging Characteristics of Task Structure to Predict the Cost of Interruption. In: Proc. of SIGCHI Conference on Human Factors in Computing Systems, pp. 741–750 (2006)
9. Lee, J., Marsella, S.C.: Nonverbal behavior generator for embodied conversational agents. In: Gratch, J., Young, M., Aylett, R.S., Ballin, D., Olivier, P. (eds.) IVA 2006. LNCS (LNAI), vol. 4133, pp. 243–255. Springer, Heidelberg (2006)
10. Kendon, A.: Some Functions of Gaze-direction in social interaction. Acta Psychologia 26, 22–63 (1967)
11. Kobayashi, K., Yamada, S.: Informing a User of Robot's Mind by Motion. In: CIRAS 2005, SS4B-3 (2005)
12. Yuasa, M., Mukawa, N., Kimura, K., Tokunaga, H., Terai, H.: An utterance attitude model in human-agent communication: from good turn-taking to better human-agent understanding. In: Proc. of CHI 2010, pp. 3919–3924 (2010)
13. Tanaka, T., Fujita, K.: Study of User Interruptibility Estimation Based on Focused Application Switching. In: CSCW 2011 (2011)
14. Kimura, K., Tanaka, T., Fujita, K.: Study of Head Motion Components related Uninterruptibility during Desk Working. Info. Processing Society of Japan 52-4 (2011) (in Japanese)

Ontological Approach to Aesthetic Feelings: A Multilingual Case of *Cutism*

Akifumi Tokosumi and Fumina Teng

Department of Value of Decision Science, Tokyo Institute of Technology,
2-12-1 Ookayama, Meguro-ku, Tokyo, 152-8552, Japan
{akt,fumina}@valdes.titech.ac.jp

Abstract. This paper outlines how the Sensibility-Aesthetics Ontology project aims to construct a comprehensive knowledge-base of sensibility-aesthetics related concepts. The target of the current paper is to examine 'kawaii' (*cutism)* in Japanese and Chinese culture. Using a combination of four methodologies (meta-analysis of dictionary definitions, quantitative analysis of linguistic usage, content analysis of voluntarily produced text containing the target concept, and structured interviews concerning the cognitive components of the target concept), cognitive components of the concept were formalized into a frame structure. The typical frame for Japanese *cutism* is: experiencer - mother, female, male; user - female; appraisal - small, weak; object - girl, baby, infant, pet, everyday things, animal, plant; subjective experience - attachment, charmed; action inclination - protect, possess, emotion share, conceal. Chinese *cutism* has similar framing terms although the 'object' component has far fewer variants than Japanese. As an application of the constructed *cutism* ontology, an analysis of the *cutism* words used in the novels of Haruki Murakami was discussed.

Keywords: Cutism, Aesthetic emotions, Ontology, Frame, Semantics.

1 Introduction

The Sensibility-Aesthetics Ontology project is collecting and analyzing lexical items and concepts in the following semantic fields: Sensory {red, soft, sweet, ...}; Sensibility {beautiful, pure, gaudy, ...}; Evaluations {good, suitable, reasonable, ...}; Values {sacred, patronizing, terrible, ...}; Attitudes {love, hate, approving, ...}; Emotions {happy, sad, angry, ...}; Volition {wishes, decisions, plans, ...}; Beliefs {believe, doubt, pray, ...}; Intention {praise, accuse, ask, ...}. For the human-computer interaction (HCI) task in general, identifying a participant's state of mind plays an essential role. Sensibility-Aesthetics Ontology is aiming to capture the most subtle and delicate states of mind to benefit the most advanced HCI projects.

The target concept of the current paper is 'kawaii' (*cutism*) in Japanese and Chinese. The concept was chosen on the basis of its popularity and idiosyncrasy in the Far East Asia culture.

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 161–164, 2011.

2 Method

The sensibility concept *cutism* is closely associated to the aesthetic emotion of *cutism*. To analyze the internal structure of sensibility-aesthetics concepts, we have been developing a synthesis of four methodologies: (a) meta-analysis of dictionary definitions; (b) quantitative analysis of linguistic usage; (c) content analysis of voluntarily produced text containing the target concept; and (d) structured interviews covering the cognitive components of the target concept. We applied the multi- methodology [1] to Japanese and Chinese *cutism*.

For the multi-methodological approach to sensibility concepts, we have developed a theoretical framework of the emotional processes at work. As shown in Fig. 1, we are employing cognitive appraisal theories of emotion.

Fig. 1. Cognitive appraisal framework for sensibility concept: A case for *cutism*

For the meta-analysis of dictionary definitions, we analyzed four Japanese dictionaries and two Chinese dictionaries. We used two methods to gauge *cutism*: a 'cause/effect' relationship and an 'object-situation-emotion-action' scenario, we used these to unpick the definition descriptions. Fig. 2 shows a case in a Japanese dictionary.

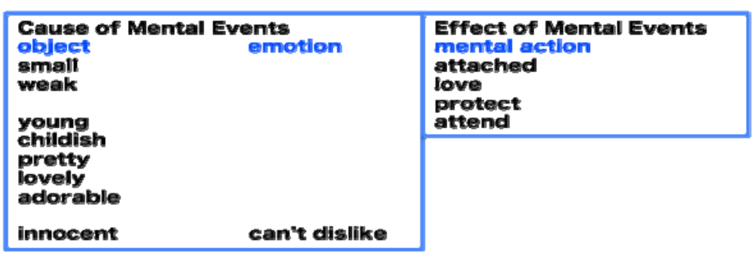


Fig. 2. A meta-analyzed slot structure for *cutism:* The case of *Daijisen* [2]

For the quantitative analysis of linguistic usage, we obtained semantically annotated word frequency data from the Case Frame Compilation corpus [3] developed at Kurohashi-Kawahara Lab, Kyoto University. The corpus is based on 470M Japanese sentences collected by a web crawler, and provides case frame (semantic deep case structure) frequency data for a given word.

The content analysis of voluntarily produced text containing *cutism* was employed to the data collected by our own web crawler and text analyzer tools [4].

The structured interviews covering the cognitive components of *cutism* starts with a question 'Please look at the word (kawaii). What do you have in your mind?'. Then the interview expands with questions related to involved situations, people, goals, issues, events, and wishes.

3 Results and Discussion

The results from the multi-methodology analysis were formalized using the cognitive appraisal framework described in the section 1. We have chosen the frame representation for the ontology. The typical frame for Japanese *cutism* is: experiencer - mother, female, male; word user - female; cognitive appraisal - small, weak; object - girl, baby, infant, pet, everyday things, animal, plant; subjective experience - attachment, charmed; action inclination - protecting, possessing, emotional sharing, concealing (Fig. 3).

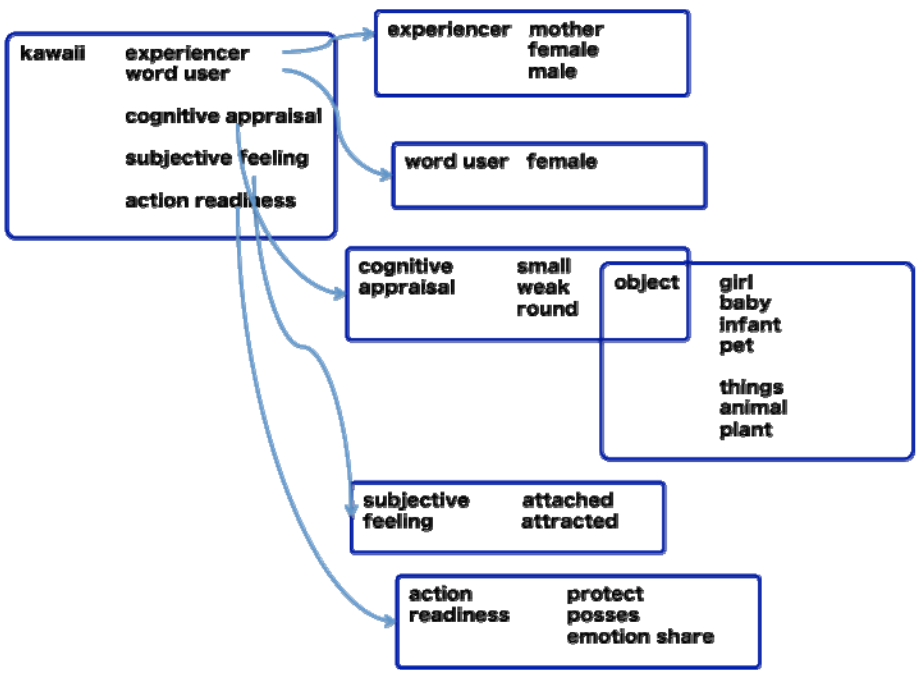


Fig. 3. A Frame representation of *cutism*: A case for Japanese

Chinese *cutism* has similar framing although the 'object' component has far fewer variants than Japanese.

As a practical application of *cutism* ontology, we analyzed how an author used the concept throughout their works over time. In Haruki Murahami's 11 novels, we identified 58 uses of the word *kawaii.* Almost all these cases are categorized as a traditional 'young lovely lady', without any deviation over the 11 works. We can infer from this that his concept of *cutism* follows a consistent pattern. Also working as a translator, Haruki Murakami [5] recently (2010) published a Japanese translation of Raymond Chandler's 'The Little Sister'. Although the work has been known as 'Kawaii Onna (cute woman)' for several decades since Shunji Shimizu's [6] translation (1959), Haruki Murakami abandoned the well-received Japanese title and adopted a simple transcription of the original title in katakana. Judging from the use of *kawaii* in his novels, we can conclude that Haruki Murakami's mental concept of *cutism* is not broad enough to include the pejorative aspects of the female character as *kawaii*.

4 Conclusion

The multi-methodology approach described in this paper proved to be useful for an analysis of the sensibility-aesthetic concept of *cutism*. Its dynamism over time and across users appropriately captures this.

Acknowledgment. This work was supported in part by Grants-in-Aid for Scientific Research 20300074.

References

1. Urata, Y., Tokosumi, A.: A Multi-Methodology Analysis of Sorrows for Emotion Ontology Construction (in Japanese). In: Third Spring Conference on Kansei Engineering, CD-ROM, p. 06 (2007)
2. Matsumura, A.: Daijisen (in Japanese). Shogakukan, Tokyo (1998)
3. Kawahara, D., Kurohashi, S.: Case Frame Compilation from the Web using High-Performance Computing. In: LREC 2006, pp.1344–1347 (2006)
4. Kawashima, T., Murai, H., Tokosumi, A.: Text Analysis Tools for Digital Humanities and Technology (in Japanese). In: 12th Annual Conference on Kansei Engineering, CD-ROM, 1C2-2 (2010)
5. Chandler, R.: The Little Sister (1949). Japanese Tr. Murakami, H. Hayakawa Shobo, Tokyo (2010)
6. Chandler, R.: The Little Sister (1949). Japanese Tr. Shimizu, S. Tokyo Sogensha, Tokyo (1959)

Constructing Phylogenetic Trees Based on Intra-group Analysis of Human Mitochondrial DNA

Ivan Vogel[1], František Zedek[2], and Pavel Očenášek[1]

[1] Faculty of Information Technology,
Brno University of Technology,
Božetěchova 2, 612 66 Brno, Czech Republic
xvogel01@stud.fit.vutbr.cz, ocenaspa@fit.vutbr.cz
http://www.fit.vutbr.cz/research/groups/nes@fit/index.php.en

[2] Department of Botany and Zoology,
Masaryk University,
Kotlářská 2, 611 37 Brno, Czech Republic
fzedek@gmail.com
http://www.sci.muni.cz/botzool/?lang=en

Abstract. This paper describes a modified algorithm for inferring phylogenetic trees based on distance techniques. The input of the algorithm consists of predefined clusters of data. It uses a usual agglomerative approach, however it involves a novel technique for distance matrix creation as the task of clustering predefined groups of human mitochondrial DNA sequences should be fulfilled.

Keywords: intra-group analysis, substitution model, position-specific clustering vector, mtDNA, population divergence, phylogenetic tree, neighbor joining.

1 Introduction

Phylogenetic tree inference is a very common method for visualising evolutionary relationships among species. Furthermore, thanks to the significant progress in the field of molecular biology, we are now able to work with organisms on the molecular level and thereby analyse nucleotide sequences, which can potentially bring us much more information about the species comparing with phenotype phylogenetic methods. The human mitochondrial genome is commonly used for studying the origins and migrations of different human populations. Mitochondrial DNA is intended for this purpose because of the relatively high mutation rate in comparison to the corresponding nuclear DNA. Moreover, forensic laboratories occasionally use an mtDNA comparison to identify human remains. In our paper, we are going to apply a novel technique of clustering of predefined clusters to identification of whole populations of human individuals.

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 165–169, 2011.

2 Problem Definition

Every distance method for constructing phylogenetic trees uses a single biological sequence as its particular input unit. That means that every leaf node in the result tree matches exactly to one input sequence. By designing our new algorithm, we work with another assumption. Let's suppose that we have a set of DNA sequences that can be classified into disjunct groups/clusters with high membership probability. We indeed assume that sequences from one group are closely related and their intra-group evolution distance is smaller in comparison to distances of sequences from other groups. In some cases there might be greater intra-group average distance which means that not every sequence of a group would be present in the same subtree of standard phylogenetic analysis. In this instance, there must be an effort made to estimate the probable position of the aggregated node with high accuracy according to the elements of the group.

The goal is to find a proper representation for every predefined group. One possible solution is to randomly choose a representative sequence for each group. Another one is to build a consensus sequence for each cluster. There is, however, a certain loss of information in both cases. We therefore present another solution using frequence analysis of predefined clusters (see subsection 2.2).

2.1 General Distance-Algorithm Template

A phylogenetic analysis of any set of biomolecular sequences based on distance metrics uses the following algorithm template:

Input: set of unaligned biomolecular sequences
Output: bootstrapped result tree

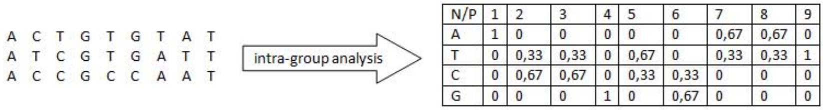
1. Application of multiple alignment on input sequences
2. Phylogenetic distance estimation and distance matrix creation
3. Application of appropriate distance method (mostly neighbor-joining [3])
4. Statistical evaluation of tree topology (mostly bootstrapping [4])

For our algorithm modification, points two and four are crucial.

2.2 Intra-group Analysis

Let's assume we have a predefined cluster. There is the need to find a proper representation of this data structure for future distance estimation among other clusters. We perform an intra-group frequence analysis of every single apriori cluster. The situation is depicted in figure 1.

We count for every single column of the group $\mathbb{X}$ (in fig.1 repesented by three sequences) so-called position-specific clustering vector (hereinafter PSCV), which contains the relative occurence of nucleotides (T, C, G, and A) in a concrete position. We thereby receive a sort of representative sequence in the form of a simple table, on which the probabilities of nucleotide occurences for every sequence position are depicted. It is straightforward that the sum of elements of every single PSCV must be 1 (as the sum of probabilities of nucleotide states in the site has to be exactly 1).



N/P	1	2	3	4	5	6	7	8	9
A	1	0	0	0	0	0	0,67	0,67	0
T	0	0,33	0,33	0	0,67	0	0,33	0,33	1
C	0	0,67	0,67	0	0,33	0,33	0	0	0
G	0	0	0	1	0	0,67	0	0	0

Fig. 1. Cluster data transformation to frequence analysis table

2.3 Distance Estimation between Two Distinct Clusters

The task in this subsection is to estimate the number of substitutions between two distinct clusters, that is, how many substitutions do we need to perform to get from one cluster to another cluster. Let's assume we have cluster $\mathbb{A}$ and cluster $\mathbb{B}$ and their PSCVs on position n, that is $v_{\mathbb{A}}[n]$ and $v_{\mathbb{B}}[n]$, respectively.

$$v_{\mathbb{A}}[n] = \begin{pmatrix} p_{\mathbb{A}}^{T} \\ p_{\mathbb{A}}^{C} \\ p_{\mathbb{A}}^{A} \\ p_{\mathbb{A}}^{G} \end{pmatrix}, v_{\mathbb{B}}[n] = \begin{pmatrix} p_{\mathbb{B}}^{T} \\ p_{\mathbb{B}}^{C} \\ p_{\mathbb{B}}^{A} \\ p_{\mathbb{B}}^{G} \end{pmatrix} \tag{1}$$

To attain the probability of substitution from nucleotide T to T (which means, both clusters contain nucleotide T at this position), we simply multiply $p_{\mathbb{A}}^{T}$ by $p_{\mathbb{B}}^{T}$, which goes for the three remaining nucleotides, as well. That is, to get probability N_n, that no substitution occurs at position n, we simply perform a dot product of $v_{\mathbb{A}}[n]$ and $v_{\mathbb{B}}[n]$. The probability of substitution at this position is therefore $N_s = 1 - N_n$.

The evolution distance between two nucleotide sequences can be estimated with the Jukes-Cantor substitution model [2] (see equation 2). $\hat{p}$ in equation 2 stands for the proportion of substitution sites to all sites in the examined sequences and is also known as p distance [5]. We extend this model combining with the previously mentioned theoretical explanations. That is, to estimate $\hat{p}$ we add N_s values of every nucleotide position together and divide it by the length of the representative sequence (number of rows in table from figure 1).

$$\hat{d} = -\frac{3}{4} ln(1 - \frac{4}{3}\hat{p}) \tag{2}$$

3 Case Study on Human mtDNA

We applied our new method to human mitochondrial DNA analysis. It consists of whole mitochondrial genomes of individuals from all over the world. Analysed sequence groups are listed on table 1.

The data come from [1] except the last two records. We additionaly obtained 4 chimpanzee mitochondrial genome sequences from [9].

The motivation was to reconstruct a phylogenetic tree of different human populations with our algorithmic solution and compare it with a relevant previously published study. We worked with the assumption that intra-group variability of a specific nation is smaller than that between different populations.

Table 1. List of processed human mitochondrial genome populations (and eventually a person who collected the collection in parentheses – if known) along with groups joined into and numbers of sequences worked with

Data	Predefined group	Number of sequences
European (Kivisild) Sardinian (Fraumene) Italian (Achilli)	Europe	215
Papua New Guinean (Ingman) Melanesian (Kivisild) Australian (Ingman)	Australia/Oceania	41
Japanese (Tanaka) Chinese (Kong)	East Asia	720
American [7]	America	5
African [8]	Africa	4

3.1 Data Preparation

The mitochondrial genome of humans consists of approx. 16 kbp. We simply selected highly polymorphic genome sites and their neighborhood, and with this received sequences of approximately 200 nucleotides in length.

4 Results and Conclusion

The phylogeny of predefined groups is shown in figure 2. Chimpanzees as an outgroup is placed at the base of the tree and followed by African human populations outward, from which are branches that subsequently correspond to populations from Europe, Australia/Oceania, the Americas, and East Asia. All the clades show strong bootstrap support (figure 2). The tree topology estimated in our analysis agrees well with the previously published phylogeny of 51 human populations based on 650,000 common, single-nucleotide polymorphism loci [6], suggesting that our algorithm may be a helpful tool for future phylogenetic analyses.

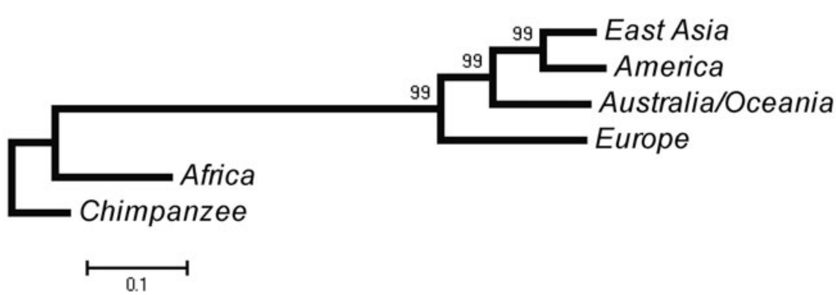


Fig. 2. Constructed phylogenetic tree with 500 bootstrap replications; the scale bar indicates p distance of reconstructed branches

Acknowledgement. This study was supported by the Czech Science Foundation (project GACR 206/09/1405), by the Ministry of Education, Youth and Sports of the Czech Republic (projects LC06073, MSM0021622416, MSM0021630528 and MSM0021630503) and projects FIT-S-11-1, FIT-S-11-2 and FEKT/FIT-S-11-2.

References

1. Ingman, M., Gyllensten, U.: mtDB: Human Mitochondrial Genome Database, a resource for population genetics and medical sciences. Nucleic Acids Research 34 (suppl. 1)
2. Jukes, T.H., Cantor, C.R.: Evolution of protein molecules. Academic Press, New York (1969)
3. Saitou, N., Nei, M.: The neighbor–joining method: a new method for reconstructing phylogenetic trees. Molecular Biology and Evolution 4, 406–425 (1987)
4. Yang, Z.: Computational Molecular Evolution. Oxford University Press, Oxford (2006)
5. Nei, M., Kumar, S.: Molecular Evolution and Phylogenetics. Oxford University Press, Oxford (2000)
6. Jun, Z.: Li, et al.: Patterns of Variation Worldwide Human Relationships Inferred from Genome-Wide. Science 319 (2008)
7. Forster, P., Harding, R., Torroni, A., Bandelt, H.J.: Origin and evolution of Native American mtDNA variation: a reappraisal. Am. J. Hum. Genet. 59(4), 935–945 (1996)
8. Herrnstadt, et al.: Reduced-Median-Network Analysis of Complete Mitochondrial DNA Coding-Region Sequences for the Major African, Asian, and European Haplogroups. Am. J. Hum. Genet. 70(5), 1152–1171 (2002)
9. Bjork, A., Liu, W., Wertheim, J.O., Hahn, B.H., Worobey, M.: Evolutionary history of chimpanzees inferred from complete mitochondrial genomes. Mol. Biol. E 28(1), 615–623 (2011)

A Qualitative Study of Similarity Measures in Event-Based Data

Katerina Vrotsou and Camilla Forsell

Linköping University, Department of Science and Technology,
Campus Norrköping, 602 74 Norrköping, Sweden
{katerina.vrotsou,camilla.forsell}@liu.se

Abstract. This paper presents an interview-based study of the definition of sequence similarity in different application areas of event-based data. The applicability of nine identified measures across these areas is investigated and discussed. The work helps highlight what are the core characteristics sought when analysing event-based data and performs a first validation of this across disciplines. The results of the study make a solid basis for follow-up evaluations of the practical applicability and usability of the similarity measures.

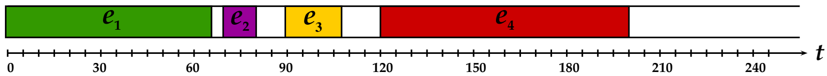
Keywords: Event-based data, event-sequences, evaluation, qualitative study, similarity measures.

1 Introduction

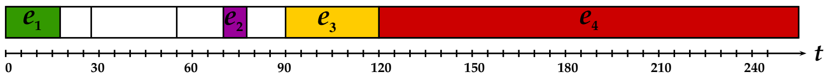
Event-based data, or event-sequence data, are encountered daily in a large range of disciplines. An event-based dataset consists of a collection of sequences of ordered events with or without a concrete notion of time [1]. This implies that the events may have an exact time-stamp associated with them or time may be relative and implicitly derived from their ordering. Examples of event-based data include medical records and procedures, time-use data, historical, biographical and career path data, internet session data, traffic incident data, process control data, and administrative process data. Even though, the data in each of these fields are of an event-based nature, they do display variations in their character. In time-use, historical, biographical and career data, for example, each recorded event has a time duration associated with it. When talking about medical event or traffic incidents the events are often regarded as instantaneous.

The process of analysing event-based data includes identifying and querying for patterns within event-sequences, comparing them and finding similarities between them. The patterns sought often have the form of shorter sequences of events, sub-sequences, which collectively exhibit a specific interesting behaviour. They may appear frequently and/or be evenly distributed, they may exhibit some sort of repetitious behaviour, or even stand out as outliers. Revealing sub-sequence similarities between data records enables their comparison, analysis and classification and can even provide a good basis for predictions.

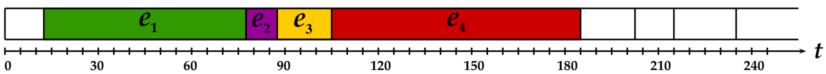
M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 170–179, 2011.



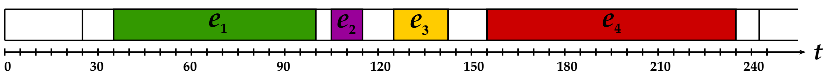
(a) Reference sequence containing the sub-sequence e_1, e_2, e_3, e_4 against which the other are compared.



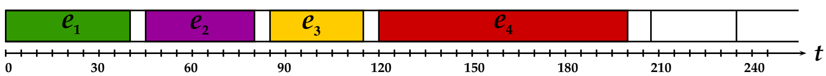
(b) The start times of the sub-sequence events are equal to the reference, their duration and intervening periods differ.



(c) The duration of each of the sub-sequence events is the same, their start times and the intervening periods differ.



(d) The sub-sequence is identical to the reference but is displaced in time.



(e) The total sub-sequence has equal duration to the reference, but the distinct events have different start times and durations.

Fig. 1. How is similarity defined? Comparison example of five event-sequences (a-e) drawn along a time-axis, t. All sequences contain the four-event sub-sequence (e_1, e_2, e_3, e_4) in different variations. Sequence (a) is the reference event-sequence against which the remaining event-sequences (b-e) are compared. The question is which sequence is more similar to the reference and why. Depending on the task and application area, different definitions of similarity are possible and hence answers may vary.

Information visualization and data-mining techniques facilitate interactive and automated analysis of event-based data by comparing and classifying individual records [1,2]. For such comparisons to be made, however, what makes sequences similar has to be determined and quantified. A set of similarity measures needs to be defined, and to do this what is characteristic, important, and/or interesting about an event-sequence has to be identified.

In previous work with time-use data, we have considered which attributes signify similarity between individuals' daily activity sequences and have identified a set of similarity measures tailored for such data [3]. In this paper we explore if and to what extent these measures have a wider applicability to event-based data. We explore four different application areas which we regard as distinct from each other and representative of several variations in the character of their data. We have performed interviews with domain experts from these areas and present and discuss the results of these interviews.

2 Motivation

Similarity as a notion can be subjective and depends largely on the characteristics of the data, the task, and the person performing the analysis. Figure 1 shows a set of event-sequences containing the same sub-sequence. The events occur in the same order but vary with respect to several measurable attributes, such as duration and start time of the sub-sequence events and duration and number of events occurring in between. The question is which event-sequences that are more similar to each other, and why. For example, when comparing work careers of individuals, the type of employment and the duration of the employment periods may be factors indicating higher similarity than the time-point during the career path when these employments started.

The motivation of this paper is to investigate how one asses similarity when presented sequences of events. What attributes define similarity in different application areas? And is there a general set of measures that can be used widely.

3 Related Work

This section will present examples of definitions of sequence similarity.

Sequence analysis is an area of high interest within computer science largely due to the extensive research in bioinformatics, which aims at identifying similarities in biological sequences based on the assumption that similarity in the structure of sequences also implies similarity in their function [4]. Techniques originally developed in this research, however, have been extended to also consider other types of event-based data. Sequence alignment [4] is such an example, in which the degree of similarity between two sequences is measured by the number of 'edit operations' (addition, deletion, substitution/move) needed to turn one sequence into the other. Each such operation carries a cost, the sum of which are referred to as edit distances, for example the Levenshtein [5] and Hamming [6] distance. Such measures define similarity based on the elements composing a sequence and the positions in which they appear. They give mostly indications of local similarity and the ordering of the elements is not directly considered. By considering sequences of elements, segments of a sequence, instead of single elements, in the similarity computation the ordering of the elements can be given importance, as in [7].

Similarity between event-sequences has been researched by Moen [8,9]. She too defines similarity based on edit operations but extends these to incorporate the occurrence time of the events [8]. This is done by assigning costs depending on the number of occurrences of an event type as well as how far an event is moved (time difference). Therefore emphasis is given to the importance of both event type and timing when comparing sequences. Furthermore, the concept of similarity based on context of events is introduced [8,9]. The context of an event is defined as the set of all event types that occur within a certain time frame before it. With this extension focus is put, not only on the event types and order of a sequence, but also on the relationships between event types and their importance in the structure of a sequence.

Matches and mismatches of events in compared sequences are considered in the sequence similarity measure presented in [10] which brings attention to the events interrupting a sequence of interest, meaning events that occur in between the matched events. This measure defines similarity in terms of event types, order and relative duration and is extended to also consider number of mismatched events, meaning the number of interrupting events [11]. Another explored measure of similarity between sequences is the frequency of occurrence of segments of consecutive elements, sub-sequences [12]. Such measures are based on the assumption that similar sequences share common sub-sequences. Similarity can thus be defined based on the number of common events and the ordering of these events, as in [13] where a similarity measure for categorical data is presented.

As evident from above, similarity between sequences can be based on many characterizing attributes which depend on the type of data as well as the objective of the task. Often, presented similarity measures are followed-up by evaluation studies performed using a certain type of data. These measures, however, are not as often evaluated across diverging application areas.

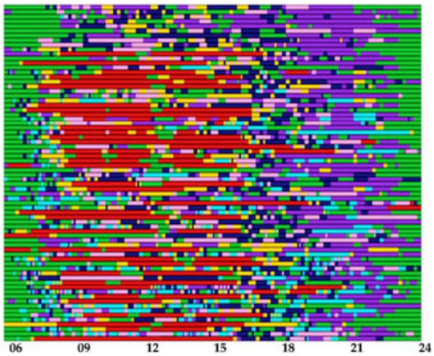
4 The Study

Nine similarity measures have been previously identified for comparing activity diary data [3]. An activity diary is composed of a sequence of activities performed during the course of the day. Within each diary, there are sequences of activities performed in order to achieve certain tasks (activity projects). For example, 'grocery shopping', 'preparing food', 'eating', and 'doing the dishes' can be regarded as parts of the activity project *'have dinner'*. Looking at how such sub-sequences are incorporated across populations (figure 2(b)) reveals similarity between individuals and facilitates the extraction of similarly behaving groups.

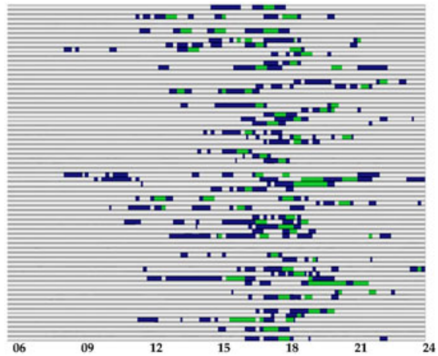
Since activity diaries belong to the larger category of event-based data, it is interesting to investigate whether the identified measures of similarity are widely applicable to this larger category. For this reason we have investigated similarity in the following four event-data types representing different application areas:

– *Time-use data*; activity diaries, biographies, careers.
– *Medical health records*; patient journal data.
– *Process control data*; air traffic control, electricity flow, paper mill industry.
– *Administrative process data*; patient administration, medical process data.

These categories are chosen because each of them represents a variation of event-based data. In time-use data events have a start time and a duration and each event starts immediately after the previous. Medical health records have a timestamp associated to each event but these do not necessarily have a duration. Also, the periods intervening events (interruptions) are of importance, especially the duration of these. Web session data and traffic incident data are also comparable to this type of data. In process control data, instead, a continuous flow is usually observed and the events have the form of abrupt changes in this flow. Administrative process and industrial process data are similar in that they are defined as



(a) Event-sequence dataset



(b) Sub-sequence highlighted

Fig. 2. Example of a sub-sequence highlighted within a dataset of event-sequence records. Similarities in how the sub-sequence is incorporated and distributed in the data become apparent. Figure (a) shows a population of women aged 25 and older and (b) the distribution of the activity project *'have dinner'* within this population. This activity project includes the activities: 'grocery shopping', 'preparing food', 'eating', and 'doing the dishes'.

processes of events that should occur and/or followed. Event-sequences reflecting how these pre-defined sequences are executed in reality can be collected in order to improve performance.

The measures that have been identified as representative for making such comparisons in activity diaries, and that are investigated here for their general applicability are listed (enumerated) below. They are concerned with the general character of each record (1-2), with the identified sub-sequence characteristics (3-6), and with the events interrupting an identified sub-sequence (7-9). In order to best explain these measures we use an example. Consider the event-sequence of figure 3. The measures will be defined by looking at how the sub-sequence 'ABC' is identified within the total event-sequence.

1. *Fragmentation* refers to the number of events composing an event-sequence; the length of a sequence. In the example of figure 3 fragmentation is 11.
2. *Variation* refers to the number of unique events composing the event-sequence; the number of distinct event types present. In the example of figure 3 the variation is 5 (A, B, C, D, E).
3. *Number of occurrences* is the number of identified occurrences of an event or a sub-sequence, meaning the number of matches found within the total event-sequence. There are 4 instances of 'ABC' in the example of figure 3.
4. *Start time* refers to the initiation time of an identified sub-sequence, meaning the start time of the first event of the sub-sequence. The start times for each of the four identified instances of 'ABC' are $< 0, 0, 18, 23 >$.
5. *Sequence length* refers to the total duration of an identified sub-sequence, from the start time of the first matched event until the end time of the last

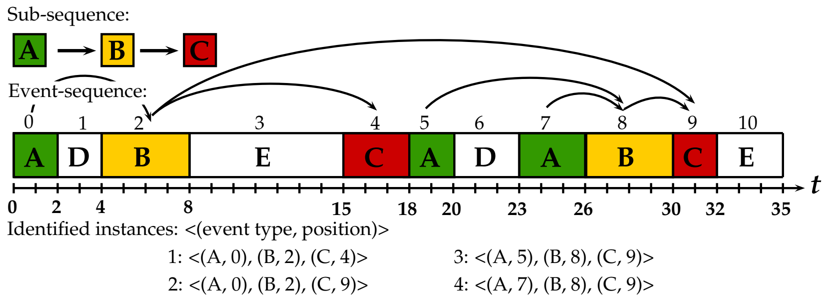


Fig. 3. Example of sub-sequence A, B, C identified within an event-sequence with 10 elements along a time-axis, t, going from $0 - 35$. Four instances are identified.

one including any possible interrupting events. The sequence length for each of the four identified instances of 'ABC in figure 3 are $< 18, 32, 14, 9 >$.

6. *Sequence event length* refers to the sum of the durations of the distinct events composing the identified sub-sequence. The sequence event length for each of the four identified instances of 'ABC' in figure 3 are $< 9, 8, 8, 9 >$.
7. *Number of interruptions* refers to the number of events interrupting an identified sub-sequence, the number of intervening events. The number of interruptions for each of the four instances of 'ABC' in figure 3 are $< 2, 7, 2, 0 >$.
8. *Length of interruptions* refers to the sum of the durations of the events interrupting an identified sub-sequence. The length of the interruptions for each of the four instances of 'ABC' in the example of figure 3 are $< 9, 24, 6, 0 >$.
9. *Type of interruptions* refers to the type of the events interrupting an identified sub-sequence, the context of the interrupting activities. In the example of figure 3 one event of type D and one of type E are interrupting the first identified instance of 'ABC'.

These nine measures cover several aspects concerning the measurement of similarity between event-sequences and we believe them to be widely applicable. They were, however, developed for a certain type of data and area. The objective with the present study has, therefore, been to explore their potential use and applicability for data analysis in other areas as well.

4.1 Method

This section describes the method applied and the results of the study .

Participants. Eight participants took part in the study, five female and three male aged between thirty and fifty-nine years. The participants were experts in the areas of time-use data (two), process-control data (two), administrative process data (two) and medical data in the form of electronic health records (two). All of them were potential users of similarity measures in their current work, however, none of them were familiar with the idea or the measures prior to the study. They received no compensation for taking part in the study.

Material and Procedure. Each participant was interviewed individually. Each session was performed as a one hour (approximately) discussion around a semi structured interview guide. This guide included twelve predefined questions and a number of potential questions used to prompt discussion. The goal was to explore the following question themes for each of the four application areas:

1. The *current process* of collection and analysis of event-based data.
2. The *attributes of interest* in the search for sequence similarity.
3. The *definition of similarity* between sequences.
4. The *applicability* of the nine proposed measures.
5. The *benefit* of the measures in terms of improving data analysis.
6. Any *additional measures* of similarity useful for data analysis.

Some days prior to the interview session each participant was provided with a written document for information and preparation. The document included a description of the background and purpose of the area and the study including descriptions of the measures and illustrations similar to figures 1 and 3. The interview session began with a joint review and discussion of this document and the interviewer provided further explanations when needed. This was done in order to facilitate the understanding of how the measures were defined and to spark the participants' reasoning and imagination on how how they could be applicable in that persons actual area of expertise.

4.2 Results

A summary of the interviews' results is presented in this section.

Current process: Two main types of analysis of processes were identified. In time-use data and in medical health records a sequence of events consists of records from a person. These are events of an arbitrary nature, number and order. The major analysis task is to compare data records of different persons, or previous data records of the same person. In process control and administrative process data a sequence is defined by a chain of predefined events. Such a sequence can be regarded as a reference and any actual outcome of a process can be compared with its ideal result.

Attributes of interest: Common for all areas was that the importance of the attributes depends on the task at hand. Occurrence and duration was recognized as important for all application areas. For time-use data and medical health records the start time of an event-sequence was critical, for example, an open wound must be sawn within a certain time frame from when it appeared. For process control and administrative processes time is also a critical attribute, not the actual start time of an event, however, but rather the duration of a sequence after it has started. See also table 1.

Definition of similarity: Here the main finding concerns the order of events as a relevant factor for defining similarity. This factor certainly has an impact on what attributes, and thus what measures, that are of interest. As mentioned above, for data in administrative processes and in process control there already exists a predefined order of events so this factor is inherent for any sequence.

Table 1. Applicability of similarity measures to the four application areas. Applicability is indicated by 'x' and inapplicability by '-'.

	time-use data	medical records	process control	administrative process
1. Fragmentation	x	x	-	-
2. Variation	x	x	-	-
3. Number of occurrences	x	x	x	x
4. Start time	x	x	-	-
5. Sequence length	x	x	x	x
6. Sequence event length	x	x/-	x	x
7. Number of interruptions	x	x/-	x	x
8. Length of interruptions	x	x	x	x
9. Type of interruptions	x	x	x	x

Hence, similarity is defined as consistency with a fixed reference sequence. For time-use data and medical health records, on the other hand, order is of great importance since different order can indicate different behaviour. The definition of similarity depends also on the definition and composition of the event-sequence. In most areas there are several ways to define the same sequence with respect to the type of considered events, their description detail, and the considered time unit. For example, in medical health records a sequence can be defined to include doctor visits, lab tests, diagnoses, and treatments, alternatively it can also include symptoms, treatments and recovery periods. In the first case the events are considered instantaneous while in the second case events have a duration which is of importance. The specification hence of the attributes that indicate similarity depends on the sequence composition.

Applicability: The applicability of the nine measures is illustrated in table 1. For time-use data the applicability of all measures was confirmed. For medical health records the applicability of two measures was variable depending on the definition of what constitutes a sequence. The first is *'Sequence event length'*. If the analysed sequence considers events instantaneous (as defined previously) then it is intrinsic that the events have no length, so the measure is inapplicable. The second is *'Number of interruptions'* which depends on the definition of the sequence and on who is in focus in the analysis. When looking at medical events as durable entities then the number, duration and type of intervening events are important while otherwise the duration that a patient has to wait between events is of interest. Also, the applicability of this measure may vary depending on who is in focus. For the patient, for example, the waiting time between events (duration of interruptions) is important, while for the doctor it may be the type of these interrupting events. In medical health records also *'Start time'* is extremely important since it dictates choice and effect of various treatments. In administrative processes, on the other hand, time is a relative attribute. Tasks have to be handled within a certain time frame after their initiation so the duration of an event or sub-sequence is of importance while *'Start time'* is irrelevant. The process control area shows the same applicability pattern as administrative processes. In general, when a sequence of events is given beforehand neither *'Fragmentation'* nor *'Variation'* are relevant. Furthermore, all aspects of interruptions are of importance in any such process.

Benefit: Much attention was focused towards optimization in all application areas. In process control and administrative processes the measures allow various comparisons of an actual outcome with a reference outcome. Hence one can detect what and where in a sequence deviations and anomalies from that idealized result are present. Here the measures can be used as tools for reflection to optimize behaviour, procedures and actions to reach a desired outcome. For time-use data and medical health records the main benefit was identified as an enhanced ability to find, identify, cluster and compare important occurrences, patterns and trends of events or sequence leading to improved analysis. Also, all participants expressed that the measures have a wide use when it comes to mining and fusion of statistics e.g., frequencies, durations and other types of various attributes of similarity.

Additional measures: In time-use data and medical health records it is interesting to compare an identified sequence of events as a whole, whereas in process data it is more important to study the characteristics of the distinct events composing the identified sequence, such as when an event starts and how long it lasts. As a result, a new measure was identified as *'Event start'* that will measure and compare the starting point of each distinct event in a sequence. This start can either be an exact time point, or the starting point in relation to the beginning of the sequence; the second day after the initiation of a task, for example. The measure is applicable to all considered areas.

5 Discussion and Conclusions

This paper is part of a larger effort to develop similarity measures to support analysis of event-based data. We have validated nine measures for time-use data and shown that they are general enough to also be applicable to other areas. The results, however, highlighted the need for careful definition of what constitutes an event-sequence since there are several possible options for each area and the applicability of the measures may vary accordingly. A summarizing conclusion is that event-based data often represent two categories of sequences.

1. *Either things happen*; events or incidents occur, activities are performed, sequences of choices are documented, and then one looks at how, when and why they happen. In general, data that represent events that are performed 'freely' fall under this category, for example activity diaries, electronic health records, incident data, web sessions and transaction data over time.
2. *Or things should happen*; sequences of events representing a process are predefined and should be followed to perform a certain task, and then one examines how close the actual outcome is to this optimally defined sequence. Data that represent the turn out of a predefined sequential process, may that be administrative or industrial, fall under this category.

For both categories, order, frequency, time and interruptions are all relevant attributes in their different forms since they are intrinsic characteristics of a sequence. Measures such as fragmentation and variation are applicable to the first category of data but not to the second category where event-sequences are predefined. Also start time is irrelevant in this case as a measure since it is the total duration of a process that is important, not when the process was initiated.

Further, we have identified an additional attribute to consider when defining similarity, namely the starting point of each event in a sequence. This resulted in the new measure *'Event start'* which is applicable to all areas.

To conclude, the study has confirmed the theoretical applicability of the measures with respect to what an analyst is interested in when comparing event-sequences. These results will provide solid basis for forthcoming development of the measures and for assessment of their use and usability in practice with quantitative evaluations.

References

1. Han, J., Kamber, M.: Data mining. Concepts and techniques. Morgan Kaufmann, San Francisco (2006)
2. Tan, P.-N., Steinbach, M., Kumar, V.: Introduction to Data Mining. Addison-Wesley Longman Publishing Co., Inc., Boston (2005)
3. Vrotsou, K.: Everyday mining: Exploring sequences in event-based data. PhD thesis, Linköping University (2010)
4. Durbin, R., Eddy, S.R., Krogh, A., Mitchison, G.: Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids. Cambridge University Press, Cambridge (1998)
5. Levenshtein, V.I.: Binary codes capable of correcting deletions, insertions, and reversals. Soviet Physics Doklady 10, 707–710 (1966)
6. Hamming, R.W.: Error detecting and error correcting codes. Bell System Technical Journal 26(2), 147–160 (1950)
7. Ergun, F., Muthukrishnan, S., Sahinalp, S.C.: Comparing Sequences with Segment Rearrangements. In: Proceedings of Foundations of Software Technology and Theoretical Computer Science, pp. 183–194. Springer, Berlin (2003)
8. Moen, P.: Attribute, Event Sequence, and Event Type Similarity Notions for Data Mining. PhD thesis, Dept. of Computer Science, University of Helsinki (2000)
9. Mannila, H., Moen, P.: Similarity between Event Types in Sequences. In: DaWaK 1999: Proc. of the First International Conference on Data Warehousing and Knowledge Discovery, Florence, Italy, pp. 271–280. Springer, Heidelberg (1999)
10. Wongsuphasawat, K., Shneiderman, B.: Finding comparable temporal categorical records: A similarity measure with an interactive visualization. In: IEEE Symposium on Visual Analytics Science and Technology, pp. 27–34 (2009)
11. Wongsuphasawat, K., Plaisant, C., Shneiderman, B.: Querying Timestamped Event Sequences by Exact Search or Similarity-based Search: Design and Empirical Evaluation (2010)
12. Pearson, W.R., Lipman, D.J.: Improved tools for biological sequence comparison.. Proc. of the National Academy of Sciences of the USA 85(8), 2444–2448 (1988)
13. Gómez-Alonso, C., Valls, A.: A Similarity Measure for Sequences of Categorical Data Based on the Ordering of Common Elements. In: Modeling Decisions for Artificial Intelligence, vol. 1, pp. 134–145. Springer, Heidelberg (2008)

Feasibility Study of Predictive Human Performance Modeling Technique in Field Activities

Naomi Yano[1], Toshiyuki Asahi[1], Shiníchi Fukuzumi[1], and Bonnie E. John[2]

[1] Human interface Technology group, Information and Media Processing Laboratories, NEC Corporation, 8916-47, Takayama-Cho, Ikoma, Nara, 630-0101, Japan
{n-yano@cb,t-asahi@bx,s-fukuzumi@aj}.jp.nec.com
[2] Human-Computer Interaction Institute, Carnegie Mellon University, 5000 Forbes Avenue Pittsburgh PA 15213, USA
bej@cs.cmu.edu

Abstract. This paper reports the results of usability evaluation based on the predictive human performance models applied to any products, and introduces application of a tool for predicting operational time to an IP phone and an electronic health record system UI consulting. We assessed effectiveness of this tool using efficiency estimation technique and extracted the practical the practical problems. In the human interface (HI) consultation process, as the tool predicts execution time of the current version and of the improved version, the improvement effect could be assessed. For an IP phone, we created the modified user interface designs from the point view of operational efficiency, so that we could indicate effectiveness of this tool by comparing task execution time. For an electronic health records, however, it is difficult to "directly" verify the effectiveness of the modified user interface designs from an efficiency standpoint. Through an evaluation scenario, the tool provided data that is necessary for assessment of improvement in this case.

Keywords: Usability Evaluation, Predictive human performance model, Consultation, Efficiency, Understandability.

1 Background and Objectives

Usability evaluation methods that use predictive human performance models are easy to deal with because they can evaluate without subject or prototype [1][2]. Such methods have several advantages over traditional usability testing methods: (1) low time and effort cost, (2) providing data about human performance early in the development life-cycle, (3) being used to generate quantitative usability metrics for proposed designs.

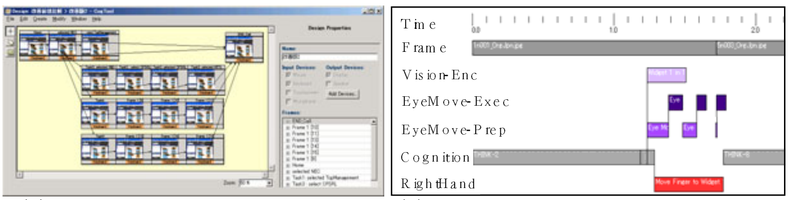
Predictive human performance modeling techniques have been investigated for many years and their feasibility regarding such factors as prediction accuracy or data input cost has been validated through many studies [3]. However, unexpected problems may still remain when they are applied to actual field activities, e.g., HI evaluation/improvement. Thus, the authors tried to apply a predictive human

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 180–189, 2011.

performance modeling technique to software products and to verify its feasibility in order to develop a usability guideline.

2 Tool for Predicting Operational Time

We selected CogTool as an example of predictive human performance modeling techniques. It is a kind of HI prototyping tool that can predict the operation time of a skilled expert in relation to a target task on the order of seconds [4]. Because HI operation times can be predicted using this tool, the operational efficiency of multiple design proposals can be compared. For this reason, it can be applied for benchmarking design proposals in the early product development stages. The design proposal is represented with a storyboard, and the correct operational procedures for executing a certain task are entered in the form of manipulating HI parts on the screen.



(a) Entry screen of storyboard (b) Visualization of the prediction

Fig. 1. Input and output of CogTool: (a) Users entry screen of storyboard on this screen. The storyboard is represented by screen images, widgets and screen transition. (b) This screen shows the computational process of the predicted time for human performance. Time runs from left to right and the widths of the boxes are proportional to the time they take to execute. The rows of boxes are different types of "operators (perceptual, cognitive or motor process, etc)" that happen in the course of performing these Tasks.

3 Methods

The tool is applied in HI consultation to assess the usefulness of a predictive human performance modeling technique. Because it can predict task execution time, we think the tool could be used in two major ways for consultation: (1) to assess the operational efficiency of multiple design proposals, and (2) to show the improvement effect of the modified user interface design. Here, we used the tool in the latter way.

3.1 Procedural Steps

We are commissioned to improve the user interface of a current system and propose a modified user interface design. This tool is applied in two consultation situations, one is virtually planned in the laboratory and the other is an actual one which has clients (developers) and their product. The procedural steps are as follows.

(1) Extract problems of the current system version by a heuristic evaluation.
(2) Focus on a problem.

(3) Write down a modified user interface design to solve the problem.
(4) Create a wire-frame image based on the hand-written modified user interface design.
(5) Apply CogTool to the current and improved system versions.
(6) Analyze the application results and assess the improvement effect.
(7) Present the results to clients. (for the real consultation case).

4 Apply CogTool to Systems

CogTool was applied to an IP phone system and an electronic health record system, which are target consulting systems. Former has efficiency problems and the latter has understandability problems. Modified user interface designs for solving them were created. Although the current implementation of CogTool only predicts performance time after a user becomes skilled (i.e., training or practice have eliminated understandability problems), and thus, we did not expect it to predict understandability issues, we applied it to both systems to see if it could show the efficacy of modified user interface designs. Analysis of these cases will enable testing of the practicality of the tool in two separate cases: assessing the improvement effect for efficiency and assessing the improvement effect for understandability. Described below are the cases of applying the tool to these systems.

4.1 Case (1) IP Phone System

This tool was applied in HI evaluation/improvement processes in which it is assumed that the tool is used in consultation for an IP phone system. Potential practical problems were discussed with two usability engineers. Because it was made clear that the operation of outgoing call needs many steps through the use of a heuristic evaluation, a modified user interface design was improved for efficiency. We set up four tasks and predicted task execution time by the tool. The predicted time of the improved version was 2.5 - 5.4 seconds shorter than the current version for three tasks and was 0.1 seconds longer than the current version for one task. We assumed the results should be enough to show the advantages of modified user interfaces objectively even in the actual consultation.

4.2 Case (2) Electronic Health Record System

The tool is applied in HI consultation for an electronic health record system. The details of the procedural steps are described in the following subsections.

Problems with Current System and Modified User Interface Designs

Consulting workers evaluated the current electronic health record system through the use of a heuristic evaluation and extracted problems. Described below are “Text box operations” and “Item selection”, which were a part of the problems and the modified user interface designs to solve them.

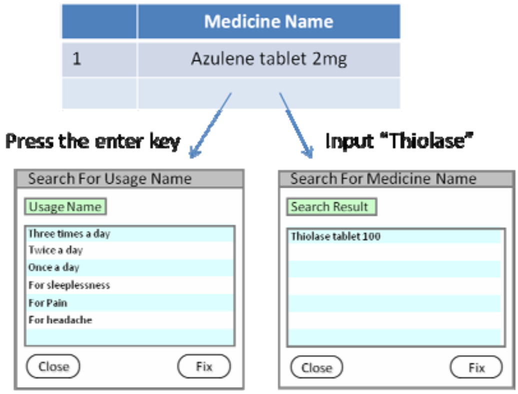
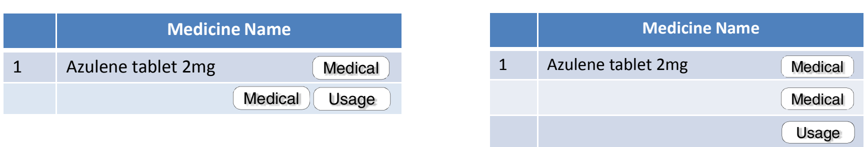


Fig. 2. Problem of "Text box operations"

Problem (1): Text Box Operations

[Problems] Users enter a medicine name in a text box and a new text box appears under it. The new text box branches the process in the input state. In the case where a user presses the Enter key without an input string, the usage search screen appears. In the case where a user enters a string, the medicine name search screen appears. For this reason, users can't predict the operational procedure visually.

[Modified user interface designs]
(1) Modified user interface design 1: The "medicine" and "usage" fields were put on the text box. This made the operation procedure clear.
(2) Modified user interface design 2: In addition to improvement (1), a text box for usage was always displayed.



(a) Design 1 (b) Design 2

Fig. 3. Modified user interface designs for the "Text box operation" problem

Problem (2): Item Selection

[Problems] Operation becomes onerous because a new window appears each time an item is narrowed down. Because the main screen hides behind many sub screen windows, a user can't see his operational behavior on the main screen. In general, a user clicks an item and presses the fix button and the selected result is reflected. But the selected result is reflected on the main screen by a single click in this interface. If a user makes a mistake in selecting an item, recovery requires many steps (he must look back at the main screen and delete the item).

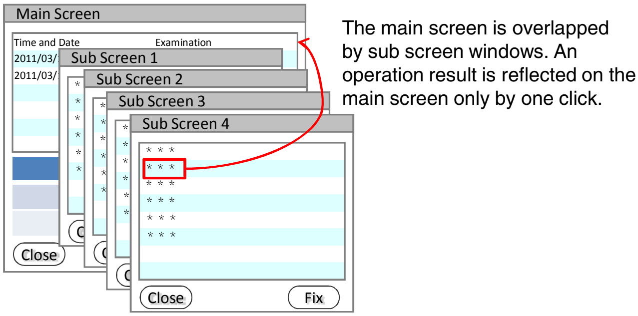


Fig. 4. Problem of "Item selection". The main screen is overlapped by sub screen windows. An operation result is reflected on the main screen only by one click.

[Modified user interface designs]

(1) Modified user interface design 1 (Single selection, depth of data is fixed):
Items are narrowed down in one window. In the last step, the "Fix" button is clicked to finalize the operation.

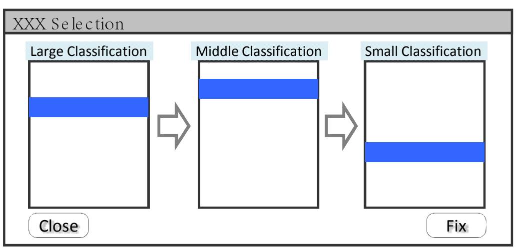


Fig. 5. Modified user interface design 1

(2) Modified user interface design 2 (Single selection, depth of data is not fixed):

・Items are narrowed down on the same screen.

・Breadcrumb navigation shows the current location.

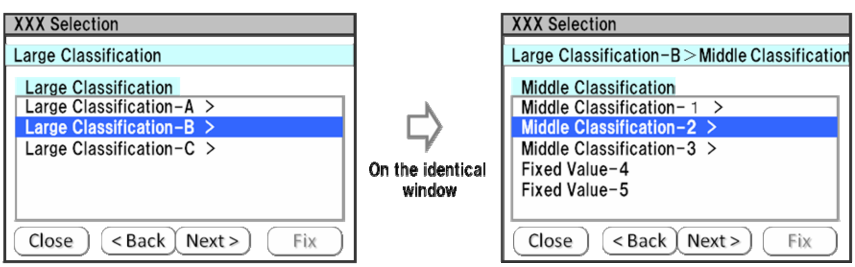


Fig. 6. Modified user interface design 2

(3) Modified user interface design 3 (Multiple selection, depth of data is not fixed):

- In addition to improvement (2), the screen has a list of selected items. The users can delete an item with the "Delete" button.
- The users push the fix button in the last step and selected items only are reflected on the main screen

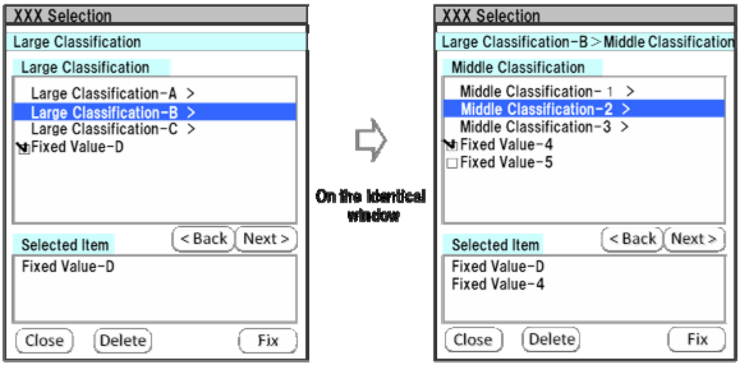


Fig. 7. Modified user interface design 3

Task Selection

The problem of "Text box operations" was that users couldn't understand the operational procedure visually. The "medicine" and "usage" buttons were put in the text box in the modified user interface design. This made the operation procedure clear. In this case, because operational objects were added, it was expected that prediction time of the improved version would be longer than the current version.

The problem with "Item selection" was that users couldn't look at their operational behavior on the main screen, and the selected result was reflected on the main screen by a single click. In the modified user interface design, items are narrowed down in one window. Clicking the "Fix" button finalized the result in the last step.

In this case, it was expected that prediction time of the improved version would be as long as the current version for the correct step. However, it was expected that prediction time of the improved version would be shorter than the current version in the recovery step. Thus, it was decided that the improvement effect would be assessed with not only correct operation steps but also with recovery operation steps for operation mistakes. We selected tasks that follow these two scenarios and applied the tool to predict task execution time. "Text box operations" has one task and "Item selection" has three tasks (correct step and two recovery steps).

Result of Applying the Tool

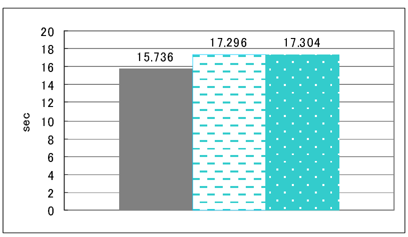
[Prediction for task execution time]

Prediction for execution time of the current version and the improved version were as follows (See Figure 8).

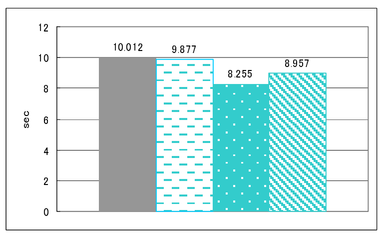
(1) Text box operations
The prediction time of the current version was the shortest. Both of the improved versions were 1.5 seconds longer than the current version.
(2) Item selection (Correct step)
The prediction times of improved versions 1-3 were shorter than the current version. The time of improved version 2 was the shortest and 1.7 seconds shorter than the current version.
(3) Item selection (Error-recovery step 1: A user executes the task with recovery operation because he made a mistake at the middle step.)
The prediction time of the current version was the shortest. The times of improved versions 1, 2 and 3 were shorter than the current version. The shorter of them was improved version 1 and 4.4 seconds shorter.
(4) Item selection (Error-recovery step 2: A user executes the task with recovery operation because he made mistake at the last step.)

The times of improved versions 1 and 3 were much shorter than the current version. The time of improved version 2 was 2.1 seconds shorter than the current version. The shorter of them was improved version 3 and 16 seconds shorter.

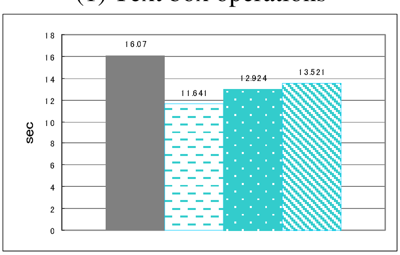
[Analysis of prediction time]
To analyze the cause of the prediction time differences, the timeline that shows the computational process of the predictions for human performance was checked. Described below is only the result of analysis for "Text box operations."



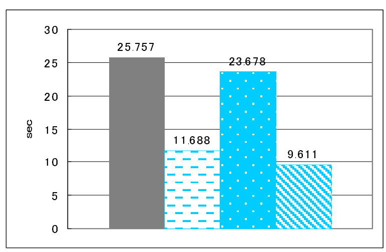
(1) Text box operations



(2) Item selection(correct step)



(3) Item selection (error step 1)



(4) Item selection (error step 2)

: Current Version, Improved Ver.1,
: Improved Ver.2, Improved Ver.3

Fig. 8. Prediction time for each task

Figure 9 shows that procedural step ④ causes the temporal difference between the current version and improved version, that is, the difference between the procedure to only press the enter key on a text box in the current version and the procedure to move the cursor to the "Usage" button with a mouse and click it in the improved version. By contrast, location change of the "Together" check box in the improved version didn't cause temporal difference between the current version and improved version.

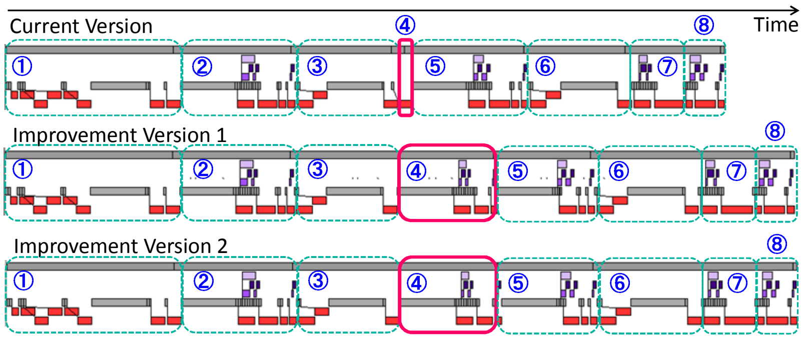


Fig. 9. Timeline of predictions in "Text box operations" task: Box ① - ⑧ represent procedural step. The step ④ causes the temporal difference of prediction time.

Evaluation Result of Modified User Interface Design

[Enter in text box]
The modified user interface designs made the procedure more easily understandable and its execution time was 1.5 seconds longer than the current version. Thus, we explained "the modified user interface designs were more understandable" and "the degradation of efficiency was 1.5 seconds" to clients and gave them quantitative metrics for UI selection.

[Item selection]
In the correct step, there is not much difference of prediction time between the current version and the improved version. In the recovery step, however, the times of the improved versions were much shorter than the current version. This is because users can look at selected items before finalizing them or go back to the last state if they make a mistake in the improved versions. It failed to consider the recovery step as it attached more importance to operational efficiency in the current version.

Because the time of the improved version was 0.1 – 0.7 seconds shorter than the current version in the correct step and was also 6 – 16 seconds shorter in the recovery steps, it was found that efficiency of the improved version was higher than the current version in not only correct operation steps but also recovery operation steps for operation mistakes.

5 Lessons Learned and Future Work

[IP phone]
Prediction time difference between the current version and the improved version showed the improvement effect of the modified user interface designs because they were made to improve efficiency.

[Electronic health record]

- We made sure that task execution time was effective as a metric that shows the improvement effect of the modified user interface design because it is easily understandable for clients.
- When the improvement is made for other than efficiency attributes (e.g.. understandability), it is difficult to "directly" verify the effectiveness of the modified user interface designs from an efficiency standpoint. For this case, however, the tool also gave clients data that is necessary for assessment of the improvement effect.

We obtained the following knowledge from clients or usability engineers.

- In this paper, the tool only shows that efficiency does not go down, and does not show the improvement effect on understandability directly. It would appear that to evaluate trade-offs between decreased efficiency and increased understandability, a tool that could predict understandability would be a useful addition to usability evaluations.
- Besides execution time, it would appear that the number of clicks with a mouse, necessity of shifting between mouse and keyboard, and distance of moving the cursor would be useful for usability evaluation. These could be computed using outputs of the tool. We plan to use these attributes in usability evaluation.

In this paper, we explained that “the modified user interface designs were more understandable” and that “the degradation of efficiency was 1.5 seconds” to clients and gave them quantitative metrics for UI selection. However, we could not evaluate the improvement effect on understandability directly. Since our results showed predictions of decreased efficiency with designs that we hoped improved understandability, a tool that can predict understandability would be useful. CogTool-Explorer [5] is a research version of CogTool that uses Information Foraging Theory [6] to predict understandability. In addition, we hope develop a technique that can estimate cognitive load or its impact on user performance and to expand the applicable scope of the tool using it. We are also planning to accelerate putting this tool in practical use by implementing it into a user interface development system.

References

1. Chi, E.H., et al.: The bloodhound project: automating discovery of web usability issues using the InfoScent simulator. In: Conference on Human Factors in Computing Systems, pp. 505–512 (2003)

2. Blackmon, M.H., et al.: Tool for accurately predicting website navigation problems, non-problems, problem severity, and effectiveness of repairs. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 31–40 (2005)
3. Carroll, J.M.: HCI Models, Theories, and Frameworks: Toward a multidisciplinary science. Morgan Kaufmann Pub., San Francisco (2003)
4. John, B.E., et al.: Predictive human performance modeling made easy. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 455–462 (2004)
5. Teo, L., John, B.E.: Towards Predicting User Interaction with CogTool-Explorer. In: Proceedings of the Human Factors and Ergonomics Society 52nd Annual Meeting, New York City, New York, September 22-26, pp. 950–954 (2008)
6. Pirolli, P.: Information Foraging Theory: Adaptive Interaction with Information. Oxford University Press, USA (2007)

Surprise Generator for Virtual KANSEI Based on Human Surprise Characteristics

Masaki Zenkoyoh and Ken Tomiyama

Chiba Institute of Technology, 2-17-1 Tsudanuma, Narashino-shi, Chiba 275-0016, Japan
mzenkyou@gmail.com

Abstract. This paper proposes a technique to generate emotion state of surprise and to generate motions to express this emotion for robots. Surprise emotion is known to have unique characteristics compared to other emotions such as anger and grief. It is often caused by discrepancies between prediction and actual event and is also an instant emotion. A conceptual structure of a surprise generator to generate emotion state of surprise for robots is developed based on these human surprise characteristics. The generator contains a predictor that is used to fined discrepancies between the real events and expectations.

1 Introduction

The main objective of this paper is to propose a surprise emotion generator as an appending module for our Virtual KANSEI (VK) for robots [1] [2] and to develop an essential sub-module. The VK for robots is a mechanism to bestow robots with an ability to behave emotionally. We have developed several VKs [3] [4] that are capable to generate five emotions, joy, anger, grief, fear and disgust out of six emotions that Ekman defined in [5]. Here, the sixth emotion, namely surprise, is considered and a surprise generator is proposed to append the existing VK.

We have proposed and been developing care-worker support robots to attend to the shortage of care-workers in rapidly aging Japan [2]. A care-worker support robot aims at reducing physical and mental strain on care-workers. We consider communication between a care-worker support robot and a care-worker is of essential importance for improvement of care-worker's mental condition. Recent research works on the robot-therapy using autonomous robot "Paro" show healing effects of interaction between human and robot [6]. This is a good indicator of the importance of robot-human communication.

The VK bestows a care-worker support robot capability to behave emotionally. The Virtual KANSEI of the wide sense (VKWS) consists of the KANSEI detector, KANSEI Generator, and KANSEI Expressive Regulator (See Fig. 1). The KANSEI detector detects robot partner's emotion. The KANSEI generator, also called the virtual KANSEI in a narrow sense (VKNS), generates robot emotion. KANSEI Expressive Regulator modifies robot motions that are generated by logical processing part, consisting of the Logical Estimator and Task Selector, of the robot controller according to emotion generated by the KANSEI generator.

We had developed KANSEI generators using neural network (NN), Petri-net with GA [4], and Hidden Marcov Model [3] methods. As stated above, the KANSEI

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 190–198, 2011.

generator can generate all but surprise emotion of six emotions defined by Ekman. Surprise was not included because of its unique characteristics compared to other emotions [3][4] and that makes it difficult to treat surprise using the same framework as other emotions. Here, we propose a technique of generating surprise emotion state of a robot that can be integrated into the existing VK.

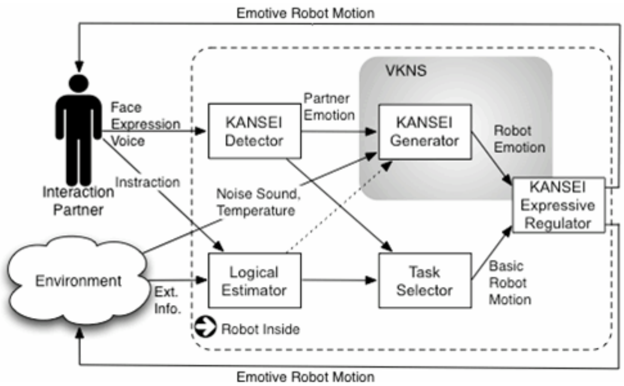


Fig. 1. Conceptual Structure of Virtual KANSEI for robots (from [3])

2 Surprise Emotion

Surprise is an important emotion defined as one of the basic emotions by many researchers, psychologist such as Ekman, philosopher such as Descarte, and so on. Following characteristics of surprise [5] make it difficult for the KANSEI generator to generate surprise emotion state for robots similarly to other emotions.

- **Occurrence factor:** Human emotions other than surprise are caused by swing of environmental condition and internal state of human. As for the surprise emotion, discrepancies between predictions and actual events often cause this emotion. Almost all discrepancies can cause surprise emotion. In addition, it is also a significant characteristic of the surprise emotion that recollection of an event that caused surprise before cannot cause surprise.
- **Time characteristics:** Surprise is an instant emotion. Surprise emotion seems to occur and disappear spontaneously at a much faster speed than all other emotions and cannot continue for an extended time. Physical expressions of surprise emotion are also shorter than other emotions [7] [8].
- **Strength of surprise emotion:** Discrepancies between predictions and actual events determine the strength of the surprise emotion. The bigger the error between the prediction and the actual events is, the bigger the strength of surprise.

Existing KANSEI generators that generate all basic emotions using the same framework are unfit for generating the surprise emotion state for robots. We propose the surprise generator as a new and stand-alone mechanism to generate the surprise emotion state for robots based on the above characteristics and as an appending mechanism to the existing KANSEI generator as shown in Fig. 2.

3 Surprise Generator

The surprise generator appended to the existing virtual KANSEI as shown in Fig.2 generates the surprise emotion state for robots. It is noted that physical expressions of surprise emotion should be limited by KANSEI Expressive Regulator for prevention of mistakes caused by excessive motions from surprise.

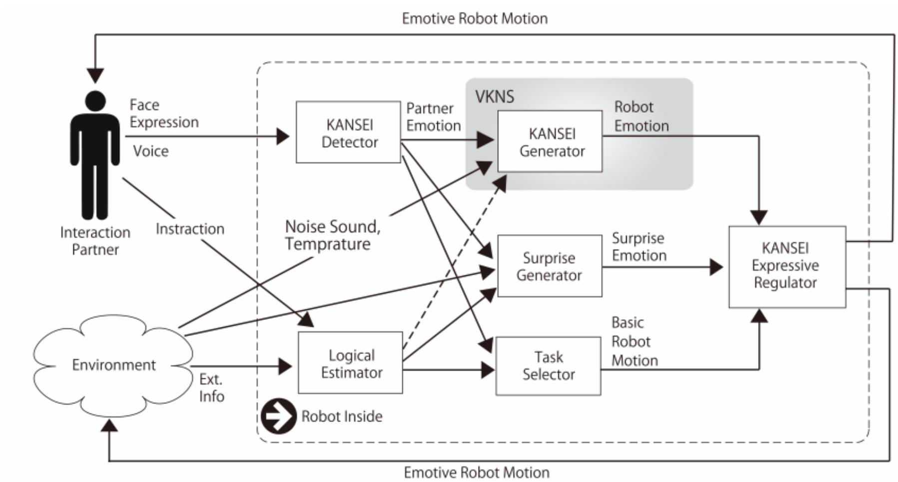


Fig. 2. Conceptual Structure of Virtual KANSEI augmented Surprise Generator

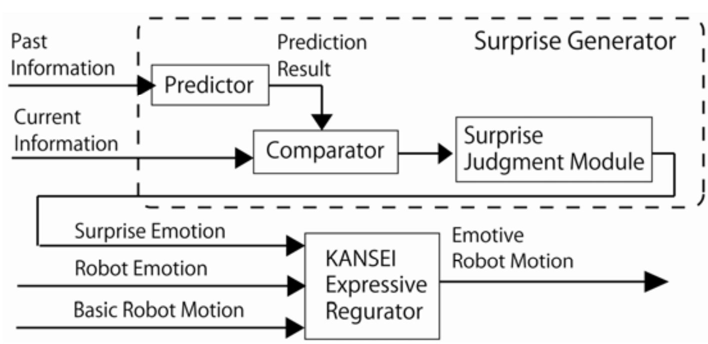


Fig. 3. Conceptual Structure of surprise generator

The proposed conceptual structure of the surprise generator as shown in Fig. 3, to become a component of the virtual KANSEI, consists of “Predictor,” “Comparator” and “Surprise Judgment Module.” Predictor predicts various information of robot partner and its environment and sends the predicted information to Comparator. Comparator compares outputs of Predictor with actual measured values. Surprise Judgment Module determines parameters such as strength and duration based on the input from Comparator and sends those parameters as the output of the surprise generator to KANSEI Expressive Regulator. The virtual KANSEI of the wide sense expresses surprise emotion based on this information.

As an initial attempt in constructing a surprise generator, we have developed Predictor that predicts sound information and human emotion transition.

3.1 Prediction of Sound Level

People are often startled by a sudden noise. We observed shifts of sound for three sound samples; human voice sound, explosive sound, and crescendo sound whose magnitude increases gradually. The following formula, Eq. (1), is proposed for predicting the sound level where s(k) is the value of sound data at time k and $s_p(k+1)$ is the predicted sound level calculated at time k. The reason for the absolute sign in this formula is to ignore the phase of the sound.

$$|s_p\,(k+1)| = |s(k)| + \{|s(k)| - |s(k-1)|\}, \qquad k = 1, 2, 3, ..., \tag{1}$$

3.2 Results and Findings of the Prediction of Sound Level

Fig. 4 is the errors between predictions calculated using Eq. (1) for three kinds of sound samples as shown in (a) for human voice, (b) explosive sound and (c) crescendo sound. Although, the explosive sound and crescendo sound can be distinguished from human voice, those two sounds have similar error levels. This result implies that the prediction method using Eq. (1) cannot be used as an indicator for sudden sounds.

The above result was obtained because Eq. (1) is essentially a differentiator and, as such, is susceptible to high frequency noise. This lead us to an adoption of a smoothing filter in the form of a simple moving average (SMA) filter. Predictor predicts three kinds of sound samples after smoothing of these sound data using the SMA filter to improve detection capability of sudden changes in sound level. Fig. 5 is the errors between predictions and actual values of the sound level that smoothed using the SMA filter. The plotted errors clearly show the difference between the explosive sound and others. This results proves that the prediction method for sound data smoothed by the SMA filter can be used as an indicator for the surprise emotion.

3.3 Prediction of Human Emotion Transition

Prediction of emotion transition of the interaction partner of a robot is realized by the KANSEI generator that was originally used to generate emotion state for robots themselves [4]. In this paper, the KANSEI generator consisting of Petri-net and Genetic Algorithm (GA) as shown in Fig. 6 is used as the KANSEI detector to predict emotion transition of the partner and assumes a role of a predictor of the partner’s emotion.

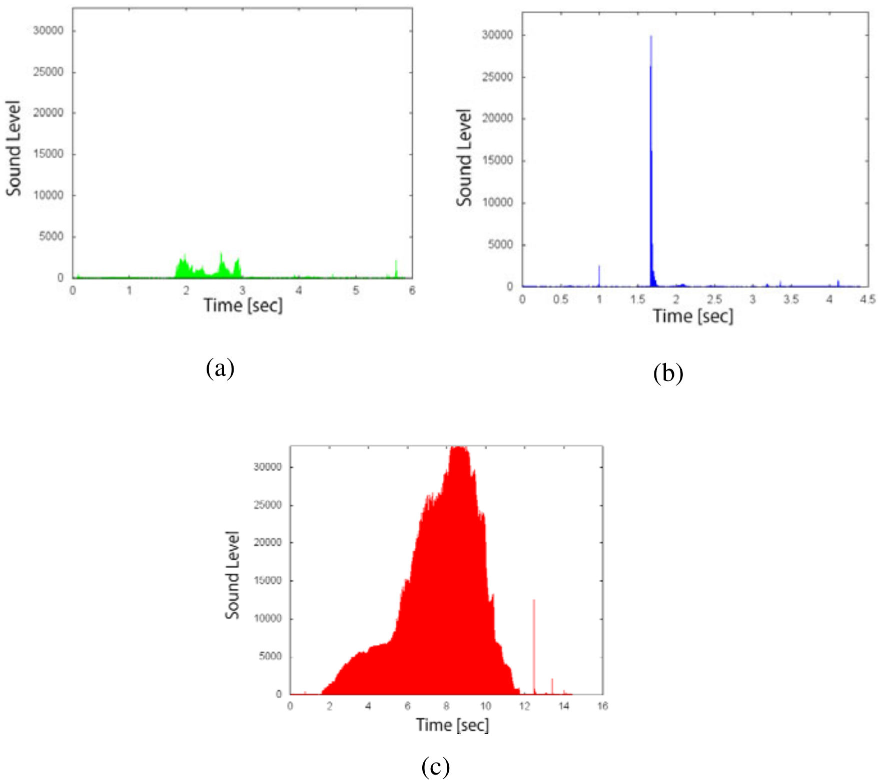


Fig. 4. The errors between the predicted sound level and the actual sound level of (a) human speaking voice, (b) explosive sound and (c) crescendo sound

The result of training is measured using the formula in Eq. (2) where the # of correct answers means the numbers that predicted emotions are consistent with the actual emotions and # of all samples means the number of total predictions.

$$\text{Evaluation Value (\%)} = \frac{\#\,of\ Correct\ Answers}{\#\,of\ All\ Samples} \tag{2}$$

Errors between the predicted emotion and the emotion detected by the KANSEI detector are quantized using the two dimensional emotion model in [1] as shown in Fig. 7.

3.4 Results and Findings of the Prediction of Human Emotion Tradition

We trained the KANSEI generator using GA and 752 teach data to predict the emotion of the interaction partner of the robot at time k+1 from the partner's emotion at time k and the output emotion of the KANSEI generator at time k-1 so as to use the KANSEI generator as the emotion predictor. GA determines the weights of arcs that are components of the Petri-net model under the conditions in Table 1.

The training result is measured using the formula in Eq. (2), and the errors between the output emotion of the emotion predictor and the training data are quantized using Euclidian distances on the two-dimensional emotion model. Evaluation value of the emotion predictor is 84.3% that means 634 output emotions out of 752 output emotions are consistent with the teach data. This result proves that the emotion predictor can predict emotion transition of the partner. In addition, the errors between the output emotion of the emotion predictor and the teach data in Fig. 8 shows that the emotion predictor can be used as an indicator for the surprise emotion.

Table 1. Parameters of GA

Number of individuals	1000
Number of elite individuals	5
Length of chromosome	228
Cross-over rate	70 %
Mutation rate	1 %
Number of generations	100000
Maximum weight of arc	100

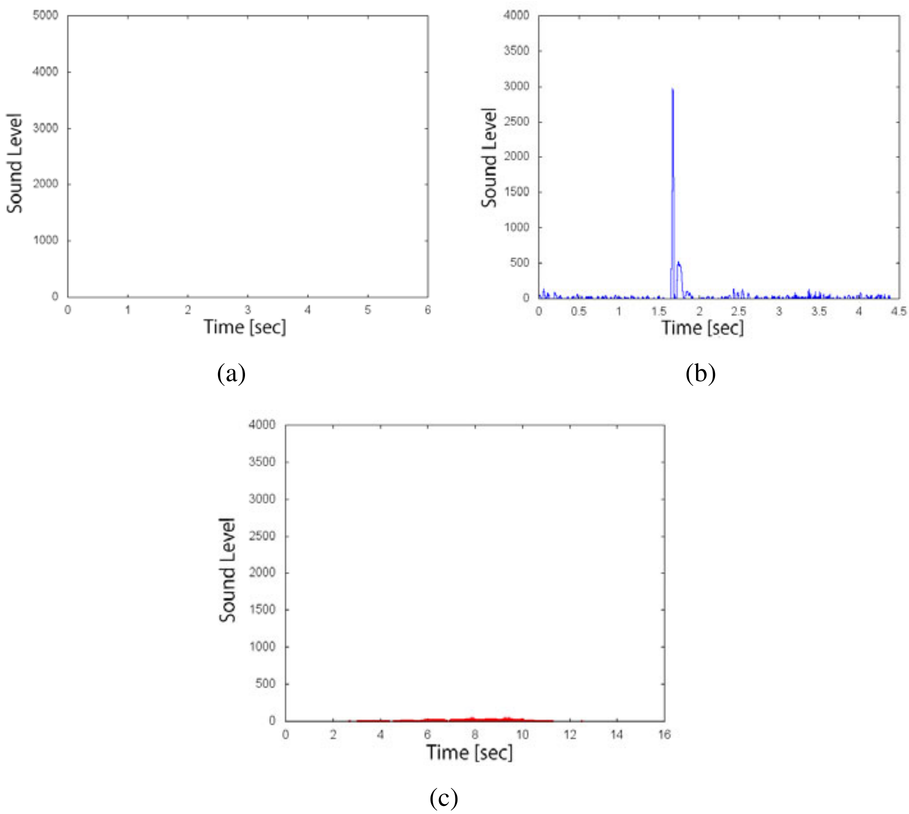


Fig. 5. The errors between the predicted sound level and the actual sound level of (a) human speaking voice, (b) explosive sound and (c) crescendo sound after applying the smoothing by the SMA filter

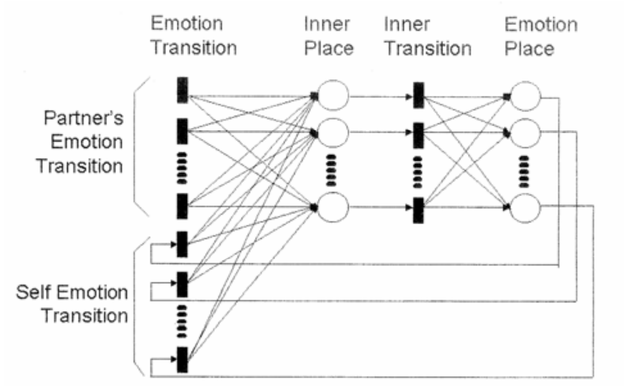


Fig. 6. The Petri-net model of the virtual KANSEI (from [5])

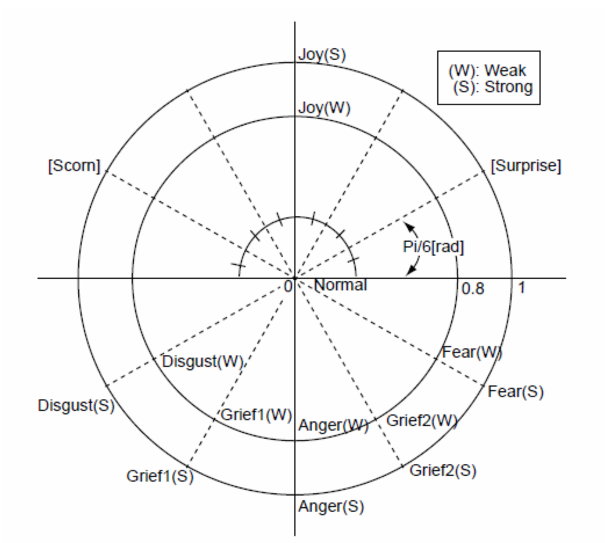


Fig. 7. Two Dimensional Emotion Model in [1]. The Euclidian distance between the two emotions, the output emotion from the surprise generator and the human emotion detected by the KANSEI detector is used as the quantized indicator of the differences of emotions.

Fig. 8. The Euclidian distances between the output emotion of the emotion predictor and the training data

4 Conclusions and Future Works

In this paper, we proposed the conceptual structure of the surprise generator based on unique characteristics of the surprise emotion. The surprise generator consists of three components; Predictor that predicts quantities of interest, Comparator that compares the prediction with measured values and Surprise Judgment Module that determines parameters such as strength and duration of surprise based on the input from Comparator. As a first step, we developed two modules for Predictor that can predict sound level and human emotion transition. The sound level predictor was shown to detect sudden burst of sound and the human emotion transition predictor was able to predict the partner's emotion at more than 80% accuracy. These results proved that these two predictors can be used as indicators of the surprise emotion.

We are interested in further developing the surprise generator. Predictor obviously should contain more modules to predict other physical phenomena that can cause the surprise emotion. Then, other components of the surprise generator need to be developed and the generator itself needs to be integrated into the virtual KANSEI of the wide sense.

References

1. Miyaji, Y., Tomiyama, K.: Virtual KANSEI for Robots in Welfare. In: IEEE/ICME International Conference on Complex Medical Engineering, pp. 1323–1326 (2007)
2. Tomiyama, K., Miyaji, Y.: Towards Realization of Care Worker Support Robot. In: Current Status of Welfare in Japan, pp. 301–329. IBUNSYA (2001) (in Japanese) ISBN4-7531-0217-3
3. Kogami, J., Miyaji, Y., Tomiyama, K.: Construction and Evaluation of a Virtual KANSEI System for Robots. KANSEI Engineering International 8(1), 83–90 (2008)
4. Miyaji, Y., Tomiyama, K.: Construction of Virtual KANSEI by Petri-net with GA and Method of Constructing Personality. In: Proceedings of the 2003 IEEE International Workshop on Robot and Human Interactive Communication, Millbrae, California, USA, pp. 391–396 (2003)

5. Ekman, P., Friesen, W.V.: Unmasking the face. A guide to recognizing emotions from facial clues Englewood Cliffs. Prentice-Hall, Englewood Cliffs (1975) (in Japanese)
6. Wada, K., Shibata, T., Tanie, K.: Robot Therapy at a Health Service Facility for the aged - Evaluation of Psychological Effects of One Year Experiment. The Society of Instrument and Control Engineers 42(4), 386–392 (2001) (in Japanese)
7. Kamachi, M., Kamachi, S., Akamatsu, S.: Does changing speed affect the recognition of facial expression? – Investigation of temporal characteristics in recognizing facial expressions using dynamic stimuli, Technical report of Institute of Electronics, Information, and Communication Engineers, HCS98-34, pp. 17–24 (October 1998) (in Japanese)
8. Kobayashi, Y., Naemura, M.: The Study of Emotion Extraction Method from Human Motion Expression - for Interaction with yourself -. Technical report of Institute of Electronics, Information, and Communication Engineers, MVE2004-76, pp. 41–46 (Mar 2004) (in Japanese)

Part III

Visualisation Techniques and Applications

Explicit Modeling and Visualization of Imperfect Information in the Context of Decision Support for Tsunami Early Warning in Indonesia

Monika Friedemann, Ulrich Raape, Sven Tessmann,
Thorsten Schoeckel, and Christian Strobl

German Aerospace Center (DLR), German Remote Sensing Data Center (DFD),
Muenchner Strasse 20, 82234 Wessling, Germany
{monika.friedemann,ulrich.raape,sven.tessmann,
thorsten.schoeckel,christian.strobl}@dlr.de

Abstract. A certainty model and appropriate visualization techniques are presented which are applied in a newly developed Decision Support System (DSS) for tsunami early warning deployed in Jakarta, Indonesia. Our decision support approach makes use of multi-sensor fusion and pre-computed tsunami scenario simulations to create situational awareness as basis for reasonable early warning. As the Indonesian coastline is prone to near-field tsunami scenarios decision making must take place under time-critical conditions based on incomplete and uncertain information. In order to reduce the probability and the consequences of a false decision, we have developed and employed a certainty model which implies a classification of imperfect information suitable for the tsunami early warning domain and the quantification of imperfect data. The model is mapped onto and supported by appropriate visual representations.

1 Introduction

The December 2004 tsunami demonstrated the need for an effective tsunami early warning system for the Indian Ocean. The work presented here is embedded in the German-Indonesian Tsunami Early Warning System (GITEWS) which has been operationally deployed in Jakarta as core of the Indonesian Tsunami Early Warning System (InaTEWS). The system combines a variety of sensor technologies such as terrestrial observation networks of seismology and geodesy, marine measuring sensors, satellite technologies, and pre-calculated tsunami scenarios used for multi-sensor online scenario selection by a Simulation System (SIM). The diverse sensor and simulation data is integrated, processed, and assessed by our Decision Support System (DSS) enabling the operator to make a concise assessment of the situation at hand and configure and disseminate the required warning messages [14].

What makes early warning for Indonesia unique and challenging is the extreme proximity of the coastline to the seismologically active Sunda Arc which initiates most of the tsunami incidents. A short time frame of merely half an hour is available for tsunami detection, warning, and evacuation increasing the operator's level of stress. Therefore, the generation of the best-possible situation awareness is crucial for a decision making by the operator at the earliest point in time.

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 201–210, 2011.

However, imperfection of information is common to early warning decision problems regarding sudden-onset types of disasters. Depending on the density of the sensor infrastructure, the types of sensors and communication networks used and the preprocessing required, sensor information will arrive in fragments, perhaps delayed, in arbitrary order and with initial uncertainties like pieces of a puzzle.

Research on the role of information imperfection for tsunami early warning is at its beginning. Academic research and projects driven by early warning centers in the Indian and Pacific Ocean focus on the tsunami propagation and inundation modeling and on error recognition and reduction in their forecast models [6, 10, 15]. Real-time tsunami modeling is in turn facing uncertainties and errors of real-time data sources which are used for making the tsunami forecast and deviations of the representing scenario to the reference data [15]. For the comparison of numerical and real data, some tsunami early warning systems provide therefore a display module where both time series are overlaid [1, 10, 15].

Indonesian operators in particular are forced to reason based on an incomplete and uncertain situational picture often without sufficient evidence from ocean sensors, in contrast to staff dealing with far-field tsunamis. The risk of late or false warnings which are expensive in terms of loss of warning credibility, economic loss, and even loss of lives is high. Our main motivation is therefore to optimize the tsunami early warning decision problem in terms of reducing late or false alarms by explicitly modeling and representing imperfect information throughout the decision support process. For this, we developed a certainty model which is used for three purposes:

1. Increase the operator's *situation awareness* (SA) and the confidence in it using information fusion and situation assessment techniques, including forecasting;
2. Based on 1., assess decision options and their consequences (costs) in order to propose the best option to the decision maker;
3. Based on 1. and 2., assess and propose the best point in time when to make the warning decision (trade-off between the pros and cons of "warn as soon as possible" and "decide as soon as enough information is available").

In the following, we will first outline the role of situation awareness in the decision support process. Afterwards, we will introduce our certainty model. We will conclude with supportive visualization techniques used throughout the process.

2 Situation Awareness and Decision Support

The main tasks of the InaTEWS DSS are to provide situation awareness to the operator and to support the early warning decision making process. To fulfill these tasks, concepts and methods of information fusion / sensor fusion and situation assessment/awareness have been applied while considering the specific requirements of human decision making under uncertainty and severe time constraints.

2.1 Concepts of Information Fusion and Situation Awareness

Two models of information fusion have been widely accepted. A functional definition is provided by the Joint Directors of Laboratories (JDL) which identify information fusion as the process of combining data to refine state estimates and predictions [8, 9, 13]. Endsley proposes another approach that addresses information fusion from a human perspective [4] which can be mapped to specific levels of the JDL model. Both approaches have been elaborated in detail in [11].

The Endsley model defines SA as *the perception of elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future* and therefore *a constantly evolving picture of the state of the environment*. Accordingly, the three main levels *perception*, *comprehension*, and *projection* have been identified which build the foundation for reasonable decision making and performance. During the remainder of the paper we will use Endsley's model which we adopted to the Tsunami Early Warning domain.

2.2 Decision Support Process

We use modeling result selections as another real-time input together with sensor information and a priori knowledge which are integrated and aggregated by different information fusion techniques to support the generation of situation awareness on all three levels. Thereupon an operator should be able to *decide* and *act* given a warning configuration for the event at hand. The cycle is repeated every time new information becomes available. Fig.1 shows the process in a schematic way.

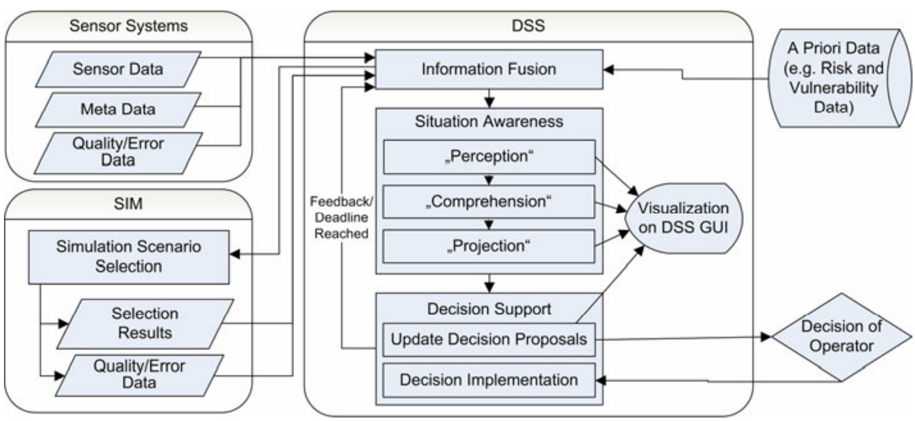


Fig. 1. The decision support core process. The system generates situation awareness based on the fusion of sensor data, modeling results, and a priori knowledge. Based on the generated SA, the system supports the operator to decide and implement the decision (*decide* and *act*).

The main task of the operator is thereby not just to warn as soon as possible risking false alarms, but rather optimize the early warning decision process under uncertainty. The early warning decision problem can be described as dynamic decision process which tries to minimize possible errors described in the following *confusion matrix*.

Both types of errors, false positive and false negative, produce costs such as casualties, economic losses and loss of confidence in the system, etc. In addition, the costs for a correct positive warning decision increase with time due to the decreasing time available for evacuation.

Table 1. Confusion matrix for the classification of decision results

Situation Assessment	Tsunami	No Tsunami
“Tsunami” Assessment	*Correct Positive*	*False Positive*
“No Tsunami” Assessment	*False Negative*	*Correct Negative*

Therefore, we extend the SA model by a planning level where the operator makes a trade-off between an immediate decision to warn or not to warn with sufficient confidence in the anticipated effects and the costs which would be caused by a wait decision to gain a clearer picture.

3 Certainty Model

We have developed a holistic model for imperfect information which we use throughout the SA generation process with the objective to promote more confident decisions. Besides providing a picture of the situation as complete and accurate as possible, we emphasize where the situation picture is imperfect. In order to serve a better comprehension of current facts and to facilitate projection into the future we take into account how data evolve over time and space.

3.1 Taxonomy for Imperfect Information

As a first step, we tried to identify all meanings of informational imperfection which are relevant to the tsunami early warning process. As we use a *more-is-better* pattern throughout our system which defines e.g. longer graphs and higher numbers as being more positive we decided to define terms of perfection instead of imperfection.

We combined the taxonomy presented in [12] with the types of uncertainty elaborated in [4]. As a result, we classify information perfection into *precision*, *certainty*, *validity*, *consistency*, and *reliability*.

Information is *imprecise* if it is incomplete, ambiguous or fuzzy whereas *certainty* relates to the knowledge of the truth of a statement. Denoting that “the earthquake has produced a tsunami of at least 1m height and I’m sure about it” is imprecise but certain whereas “the earthquake has produced a tsunami of 3.5m but I’m not sure about it” would be precise but uncertain. Both aspects can coexist but are distinct. Often any increase in one is balanced by a decrease in the other. We believe that certainty of the situational description is more crucial for a reasonable decision than uncertain precision since the latter facilitates unsubstantiated confidence.

Validity of information is determined by its timeliness. If data collection occurs continuously, information is mostly up-to-date. However, it is common to sensor technologies that data may arrive sporadically and delayed or may require preprocessing which reduces data validity. In addition, there is a connection between timeliness and certainty. Sporadic data updates introduce a degree of uncertainty as to what has changed in the interim.

Different data sources are *consistent* if they are in agreement on a phenomenon. InaTEWS DSS must be able to cope with inconsistent data. Imagine a tide gauge sensor measuring a significant wave traversal while an adjacent gauge sensor remains calm. Consistency can increase subjective confidence in the accuracy of the data. If multiple sources report conflicting information on the same situation confidence or certainty can be lowered and thus may have a negative effect on the decision process.

Uncertainty can be produced by low *reliability* of the sensors measuring the underlying phenomenon. Using our former example the second gauge sensor may have a minor reliability e.g. due to battery outage or ambient noise. Regarding tsunami detection, certain sensor types are inherently more reliable than others. The detection of an earthquake by the seismic system can *indicate* a tsunami but is less reliable regarding tsunami detection than ocean sensors which can actually prove a wave traversal.

3.2 Fitness for Purpose Analysis

Incoming data have heterogeneous attributes relating to some type of perfection. The purpose of information is to support decision making, and likewise the *quality* of information is dependent on the decisions, or the range of decisions, considered [2]. Using the classification above, we tried to identify relevant attributes provided by GITEWS input data. For instance, the list of tsunami scenarios selected by the SIM based on a given set of current sensor observations contains inherent precision and certainty in terms of list length and homogeneity. The better one or more scenarios match the measured situation distinctively the shorter the list. In addition, matching results are precise if the forecasts made by the scenarios in the list are homogenous, i.e. if all scenarios anticipate similar wave heights (estimated wave heights, EWHs) for all coastal points. In order to quantify the certainty of a tsunami scenario selection, the SIM provides a *mismatch* value as generalized distance of the scenario to the measured situation [3]. Other types of sensor observations data contain similar certainty information such as discrete error measures and confidence intervals. Regarding the reliability of a sensor each sensor is described by a number of characteristics among others a health code and a reliability value given by a sensor expert.

3.3 Quantification of Perfection

Based on the results of the fitness for purpose analysis the DSS quantifies the perfection data by a *Quality* and *Certainty* parameter which both are values in the unit interval. In the scope of the GITEWS DSS

- the Quality parameter describes the inherent perfection of information provided by *sensors* (both real sensors and SIM information before data fusion);
- the Certainty parameter describes the assessed perfection of fused information provided by multiple sensors in consideration of relations between the information.

Both parameters are currently generated using numerical methods. Quality is directly derived from perfection attributes such as the mismatch value provided by the SIM. The Certainty value is composed by fusion of individual Quality parameters and qualified by balancing factors regarding consistency and sensor reliability. The ability of sensor types to *confirm* a (tsunami) wave (e.g. ocean sensors) rather than only *indicate* the possibility that one was generated (e.g. seismology) is for instance weighted more strongly. Hence, the Certainty parameter bears - besides quantified perfection data - qualification assumptions which can be used to assess whether sufficient evidence is available for a warning decision. We differentiate levels of Certainty:

- Single-domain/single-source Certainty describes the fusion of Quality measures of different observations of one station.
- Single-domain/multi-source Certainty relates to fused Certainty information of different stations in order to assess the whole sensor system's outcome.
- Multi-domain/multi-source Certainty which we also call global Certainty is based on the fusion of Certainty information of different systems in order to assess the Certainty of the tsunami incident.

Both quantifiers can be mapped onto the SA levels. The Quality parameter is used to facilitate perception of inherent data perfection. On comprehension and projection level we make use of the Certainty value in order to quantify composite perfection, to enable comparison of sensor system outcome on an abstract level, and to make assumptions regarding potential trends. A detailed description of the mapping functions is outside the scope of this paper and will be covered by subsequent publications.

4 Visualization of Imperfect Information

Many of our design decisions are based on commonly accepted user interface principles such as *overview first, details on demand*, on the internationally standardized dialogue principles for interactive systems (ISO 9241-110) [7] and on Endsley's design principles for increasing SA [5]. Some of those SA principles address the representation of uncertainty and the increase of confidence in the context of decision support [4]. This paper concentrates on presentation techniques and patterns which reveal the perfection of the situation picture on all levels of SA.

4.1 Perception Visualization

Perception visualization provides representations of the status, attributes, and dynamics of the relevant elements as basis for the higher levels. A first SA principle is employed by explicit identification of important missing information by color, e.g. colored cell in table or colored sensor station in map. In addition, sensor reliability is presented by different types of sensor status icons. Fig. 2 shows a map extract with a gray thus offline (i.e. no data expected) buoy station next to a white thus active buoy station in tsunami mode (i.e. frequent data updates expected) and a pale-red crossed-out tide gauge station which reports invalid information (*upper left*). The table extract below shows marked missing wave height information of the gauge station in detail.

The quantifiers Quality and Certainty are mapped directly on a bar which is subdivided into fixed blocks which we believe leads to faster perception compared to numerical percentages or analog bars. Certainty bars are shown in Fig. 3.

Another SA design principle comprises the representation of information timeliness in order to assess validity of those information. We have implemented a *newest-on-top* principle by which the most current information is placed at the top of each information module. Upon update the respective data set (e.g. collection of measurements of a sensor) moves to its new position indicating risen sensor validity. A similar principle is used with the emphasis of “idle” time since the last incoming measurement in DSS time series modules. The DSS displays sea level measurements of tide gauges and buoys over time. As shown in Fig. 2 (*right*) the time range for which measurements exist is highlighted gray which we define as color representing the past. The time since the last collected measurement to a red *now*-line is marked white which makes it easier for the operator to identify which sensors produce timely thus valid information. In addition, the DSS decides which time series are displayed treating invalid sensors by default with lower priority and preferring valid sensors.

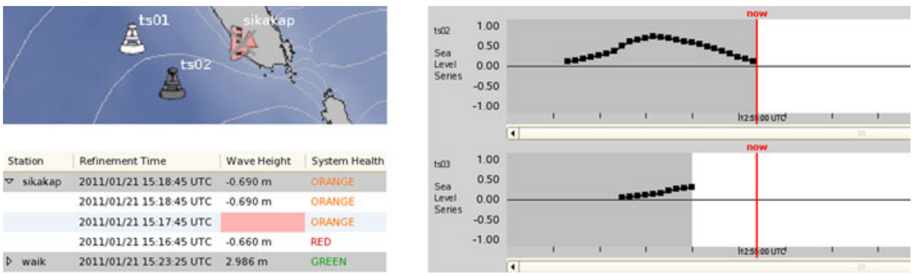


Fig. 2. Extracts of visualization modules on perception level. Map shows different types of sensor status icons, table cells with missing information are colored pale-red (*left*). Timeliness of sensors is emphasized by marking the time range since the last measurement white (*right*).

4.2 Comprehension Visualization

Comprehension includes the integration of multiple pieces of information and a determination of their relevance for the decision process in order to produce a composite picture. It takes primarily place by comparison of fused data. For this purpose, we intentionally stack the Certainty bars of all participating sensor systems (single-domain/multi-source Certainty) and disregard any details as shown in Fig. 3 (*left*). Each Certainty bar relates to the certainty of the underlying phenomenon, such as the certainty of the assessed earthquake severity (Earthquake Monitoring System) or the certainty of the assessed sea level anomaly (Tide Gauge System). By this arrangement ambiguities can be faster identified and sensor systems with incomplete or inaccurate information are discretely emphasized which conforms to another SA design principle. Based on this high-level interpretation the user can selectively explore uncertainty information in detail by analyzing the small-scale information in tables and in visualization modules to build confidence in composite data.

An example for the investigation of imperfection in composite data is provided by a graph visualization of tsunami scenario selection results shown in Fig. 3 (*right*). The filled graph shows the EWHs (y-axis) for all affected segments at the coastline (x-axis) based on an aggregation of the list of selected scenarios. The overlaid thicker graphs represent the EWHs of the single scenarios in the list. By combining single and aggregated results in one diagram the operator is able to assess the homo-/heterogeneity of the result list and thus the perfection of the matching.

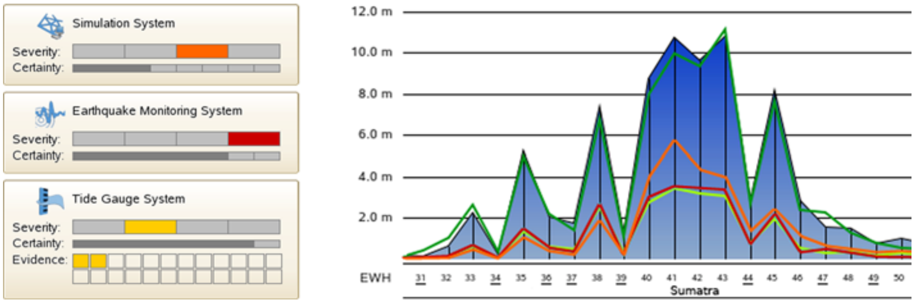


Fig. 3. Extracts of visualization modules on comprehension level. Stacked Certainty bars of different sensor systems are used for comparison on an abstract level (*left*). Heterogeneity of scenario selection result list can be explored by overlaying single scenarios of list (*right*).

Time plays an important role for the comprehension of data convergence such as the leveling-off of the earthquake magnitude value. A plot shows magnitude refinement evolution over time which helps the operator to assess the stability of the current knowledge and thus also sensor reliability. As used by other tsunami early warning systems the overlay of different time series such as simulated and real-measured sea level recordings helps understanding their consistency.

4.3 Projection and Planning Visualization

A critical part of SA is understanding how much time is available until some event occurs or some action must be taken. Time series modules and plots such as the magnitude refinement plot mentioned before help the operator to anticipate further evolution (e.g. convergence of the magnitude onto a stable value) and to schedule actions.

For the overview of the course of events and actions on a temporal scale the DSS provides a timeline module (see Fig. 4 for an extract) which is separated from left to right into past, present and future. A red *now*-line represents the present point in time. Left of it sensor measurements are displayed on the lowest lane. On the lane above ("Simulation Forecast") representations of anticipated events such as the estimated wave arrival at a sensor station are placed. Important deadlines are shown as thick blue lines. A light-blue area defines the time from now to the next important deadline.

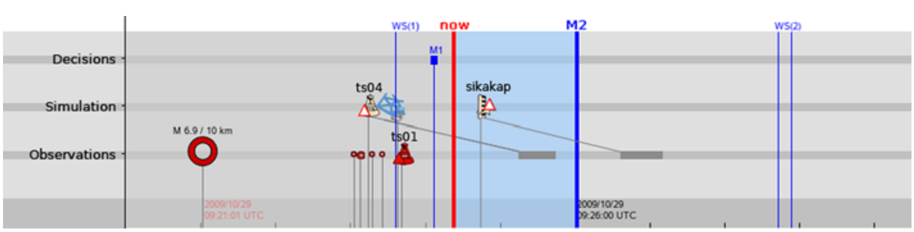


Fig. 4. Timeline visualization of sensor observations (*icons on lowest lane*), moving *now*-line (*thick red line*), important deadlines (*thick blue "M2"- line*), anticipated sensor observations and expected data arrival (*sensor icons on middle lane and connected gray blocks on lower lane*)

By examining if any buoy or gauge station is expected to be reached shortly by the wave the operator can decide whether to wait for the likely sensor data update and substantiate his knowledge or to accelerate a warning action. In the example it would make sense for the operator to wait for additional evidence from buoy *ts04* since data arrival is expected to happen before the dissemination deadline for warning message “M2” will be reached. In contrast, it could be reasonable to ignore tide gauge *sikakap* which is expected to provide recordings after the deadline.

5 Outlook

The paper describes the certainty model and related visualisation aids of the current operational InaTEWS installed in Jakarta, Indonesia. A number of additional mapping functions and visualisation tools are under development which will provide further decision support to the operator, among them probabilistic forecast models, improved time-dependent trade-off / cost functions and certainty-based consistency metrics.

Acknowledgments. GITEWS has been developed by scientists and engineers of the German Research Centre for Geosciences (GFZ, Project Lead), the Alfred Wegener Institute for Polar and Marine Research (AWI), the German Aerospace Center (DLR), the GKSS Research Centre (GKSS), the German Marine Research Consortium (KDM), the Leibniz Institute of Marine Sciences (IFM-GEOMAR), the United Nations University Institute for Environment and Human Security (UNU), the German Society for Technical Cooperation (GTZ), the Federal Institute for Geosciences and Natural Resources (BGR) as well as Indonesian and international partners. The project is funded by the German Federal Ministry for Education and Research (BMBF), Grant 03TSU01.

References

1. Annunziato, A.: The Tsunami Assessment Modelling System by the Joint Research Centre. Science of Tsunami Hazards 26(2), 70–92 (2007)
2. Arnborg, S., Brynielsson, J., Artman, H., Wallenius, K.: Information Awareness in Command and Control: Precision, Quality, Utility. In: Proceedings of the Third International Conference on Information Fusion (FUSION 2000), vol. 2, pp. pp. THB1/25–THB1/32 (2000)
3. Behrens, J.: On uncertainty and quality control in tsunami forecasting. In: Proceedings of the Indian Ocean Tsunami Modelling Symposium, Fremantle, Australia, pp. 14–18 (2010)
4. Jones, D.G., Bolté, B., Jones, D.G.: Designing for Situation Awareness. Taylor and Francis, Boca Raton (2003)
5. Friedemann, M., Raape, U., Tessmann, S., Schöckel, T., Sparwasser, N., Hochleitner, G., Strobl, C.: Decision Support Interface for Tsunami Early Warning in Indonesia – Dealing with Information Fusion, Uncertainty, and Time Pressure. In: Proceedings of IHCI 2010, Freiburg, Germany, pp. 289–294 (2010)
6. Furumoto, A.S., Tatehata, H., Morioko, C.: Japanese Tsunami Warning System. Science of Tsunami Hazards 17, 85–105 (1999)
7. DIN EN ISO 9241-110: Grundsaetze der Dialoggestaltung. Beuth Verlag, Berlin (2006)

8. Klein, L.A.: Sensor and Data Fusion: A Tool for Information Assessment and Decision Making. SPIE Society of Photo-Optical Instrumentation Engineering (2004)
9. Liggins, M.E., Hall, D.L., Llinas, L.: Handbook of Multisensor Data Fusion: Theory and Practice (Electrical Engineering & Applied Signal Processing). CRC Press, Boca Raton (2008)
10. Nayak, S., Kumar, T.S.: Indian Tsunami Warning System. In: The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Beijing, vol. 37, part B4, pp. 1501–1506 (2008)
11. Salerno, J.: Information Fusion: a High-level Architecture Overview. In: Proceedings of the Fifth International Conference on Information Fusion, vol. 1, pp. 680–686 (2002)
12. Smets, P.: Imperfect Information: Imprecision - Uncertainty. In: Uncertainty Management in Information Systems, pp. 225–254 (1996)
13. Steinberg, A. N., Bowman, C. L., White, F. E.: Revisions to the JDL Data Fusion Model. Presented at Joint NATO/IRIS Conference, Quebec City, Canada (1998)
14. Steinmetz, T., Raape, U., Teßmann, S., Strobl, C., Friedemann, M., Kukofka, T., Riedlinger, T., Mikusch, E., Dech, S.: Tsunami Early Warning and Decision Support. Natural Hazards and Earth System Sciences 10, 1839–1850 (2010)
15. Titov, V., Mofjeld, H., Newman, J., Venturato, A., González, F., Bernard, E., Eble, M.: Real-time Tsunami Forecasting: Challenges and Solutions. Natural Hazards 35, 41–58 (2005)

Kansei Stroll Map: Walking around a City Using Visualized Impressions of Streetscapes

Yuichiro Kinoshita, Satoshi Tsukanaka, and Takumi Nakama

Department of Computer Science and Media Engineering,
University of Yamanashi, Kofu, Yamanashi 400-8511, Japan
{ykinoshita,g10mk023,g09mk027}@yamanashi.ac.jp

Abstract. Information about the nature and characteristics of streets is useful to tourists. This paper introduces the *Kansei* stroll map that visualizes streetscape impressions. First, subjective evaluation experiments were conducted to investigate the impressions of 475 streetscapes in Kyoto. Using the *Kansei* structure visualization technique, proposed by the present authors, these impressions from around the city were visualized using colour and colour density and then implemented on PCs. User studies with eight participants examined its use in planning routes and walking in the city. The user studies suggest that the visualized impressions provided by the *Kansei* stroll map help users to find less-known places and streets that match their interests.

Keywords: visualization, impression, map, stroll trip, *Kansei.*

1 Introduction

Digital devices and web-based services that assist tourists have been developed and popularized for tourism. It has become common for travellers to search for information on destinations and check for efficient routes to tourist attractions using a digital map service on their mobile devices. Several studies have reported on these strategies [1, 2], which are useful for destination-based tourism. On the other hand, some travellers prefer to walk around a city without specific destinations and feel the atmosphere of districts or streets. For this way of travelling, it is more useful to get information on the characteristics or impressions of districts and streets, rather than obtaining information on the destination itself or efficient routes. Google Maps [3] provide a street view function that shows street characteristics; however, it requires the user to specify a street and it is difficult to determine the characteristics of an entire region at a time.

One method for communicating the characteristics of streets is to visualize the streetscape. The present authors have proposed a new visualization technique, called *Kansei* structure visualization [4, 5], which enables users to intuitively recognize the distribution and intensity of the impressions. *Kansei* [6] is a Japanese term that means human emotion or sensibility. This paper applies the technique to visualize the streetscape impressions in a city. This paper also implements the *Kansei* stroll map using the visualization results. The stroll map helps users to find places or streets that match their interests as they walk around a city. As a target of the visualization and implementation of the stroll map, the authors choose the city of Kyoto, Japan's old capital.

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 211–220, 2011.

2 Visualization of Streetscape Impressions in Kyoto

2.1 *Kansei* Evaluation Experiments

For the visualization step, *Kansei* evaluation experiments using the Semantic Differential (SD) method [7] were conducted to investigate the impressions of Kyoto streetscapes. This study expresses impressions in terms of bipolar adjective pairs. First, 116 adjectives related to streetscapes were collected from travel magazines, dictionaries and previous studies. By combining similar adjectives and pairing them with those having opposite meanings, ten pairs of adjectives, shown in Table 1, were obtained. For evaluation samples, 475 streetscape pictures were taken from a pedestrian perspective in the area surrounded by Karasuma St., Nijo St., Higashioji St. and Shijo St. These pictures, taken at intervals of 60–120 m, covered all the streets in the area. Although the impressions evoked by streetscapes differ with the time, weather, day and season, this study fixed the conditions at daytime, fair/cloudy weather, weekdays and early summer.

Using the adjectives and the streetscape pictures, computer-based evaluations were conducted with 10 subjects, two females and eight males. The subjects sat 210 cm away from a 65-inch LCD display. One picture was placed at the centre of the screen with a black background and a pair of adjectives with a five-point SD scale {1, 2, …, 5} was placed on the lower part. The subjects selected a suitable response for their evaluation from the scale and evaluated all the 475 sample pictures using the ten pairs of adjectives. Thus, 4750 evaluations were obtained per subject. To reduce fatigue in subjects, the evaluations were divided into ten sets and the subjects were also allowed to take breaks at any time during a set. The visualization used evaluation scores for the 475 streetscape pictures.

Table 1. Ten bipolar pairs of adjectives that express streetscape impressions

No.	Pair of adjectives	
1	modern	traditional
2	realistic	fantastic
3	uncomfortable	comfortable
4	artificial	natural
5	cheap	luxurious
6	small-scale	large-scale
7	ugly	beautiful
8	noisy	quiet
9	complicated	simple
10	unrefined	refined

2.2 Allocation of Evaluation Scores

The visualization used the average evaluation scores for all subjects with respect to each adjective pair. The scores of a streetscape picture express the evaluation for the streetscape in the picture. Because all the sample pictures featured the same perspective, i.e. approximately a 63-degree expanse, 60-m length and 30-m width from the points shown in Fig. 1(a), average evaluation scores were allocated to the area corresponding to this perspective. For obstructions in the streetscape, such as buildings, the area for the score allocation was decreased. For example, Fig. 1(b) shows the allocation of evaluation scores when buildings obstruct the view on the left.

The pictures used in the evaluations were captured at many locations, and therefore, a certain location might be included in two or more pictures and the subjects might give the pictures different scores, even though they include the same location. It is not desirable for the visualization process to receive multiple scores for one location. It is required to integrate those scores and assign a single evaluation score for the location with respect to each adjective pair. To address this problem, this study defined a parameter of the *degree of confidence*, which expresses the probability that a certain location in a picture adopts the evaluation scores for the picture. The parameter was assigned with a range of 0–1 for each location in a picture. Throughout the *Kansei* evaluation experiments, subjects paid more attention to prominent locations in a picture; therefore, locations near the places where the pictures were captured received a higher degree of confidence. On the basis of the assigned degree of confidence, the evaluation score for location l was determined as the weighted average of all evaluation scores for the pictures that included location l. When location l is included in n pictures, its evaluation score for the adjective pair k is expressed as

$$E_{lk} = \frac{\sum_{i=1}^{n} p_{il} e_{ik}}{\sum_{i=1}^{n} p_{il}} \tag{1}$$

where e_{ik} and p_{il} represent the evaluation score of picture i for the adjective pair k and the degree of confidence for the location l in picture i, respectively.

(a) no obstructions (b) obstructions on the left

Fig. 1. Areas where evaluation scores are allocated

2.3 Interpolations of Evaluation Scores

After the score allocation, locations in the sample pictures had a single evaluation score E_{lk} for each adjective pair. In contrast, evaluation scores were not allocated for locations that were not included in the sample pictures, although used in the experiments. It was however considered that these locations evoke specific impressions. This phase, therefore, allocated evaluation scores to those locations. Using all locations with evaluation scores, Delaunay triangulation divided the plane of the map into many triangles. Three vertexes of the triangles already had evaluation scores and the scores inside the triangles were linearly interpolated.

2.4 Visualization of Evaluation Scores

The visualization clarifies the distribution and intensity of the impression expressed by a single bipolar adjective pair. Score allocation and interpolation provided all map locations with an evaluation score within the range of 1–5. The score at each location was simply mapped to a colour. Figure 2 shows the colour mapping for the adjective pair *modern–traditional*. Red and blue colours represent locations with evaluation scores 1 and 5, respectively and no colour (transparent) represents locations with a score of 3. The assigned colours were overlaid on the map. Figure 3 shows the visualization of the adjective pair *modern–traditional,* where red corresponds to impressions of *modern* while blue indicates an impression of *traditional*.

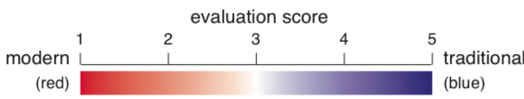


Fig. 2. Colour mapping for the evaluation scores

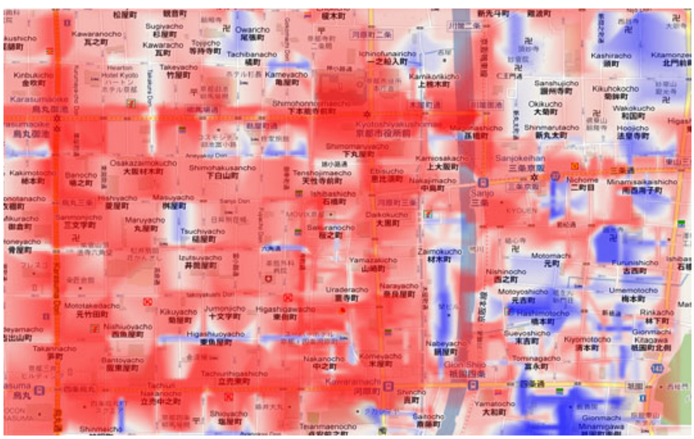


Fig. 3. Visualization result for the adjective pair *modern–traditional*

3 Implementation of the *Kansei* Stroll Map

A *Kansei* stroll map system was implemented using the ten visualization results. The screen displays a map of Kyoto with an overlay of one of the visualization results and five icons, as shown in Fig. 4. A user can pan the map by clicking and dragging. Clicking the zoom in/out icons allows users to zoom in/out the map. Clicking the switching visualization icon, located at the far right of the series of icons, displays a pop-up menu that provides a list of the adjective pairs. Users can select adjectives from the list to determine the impression to be visualized on the map. In addition to basic map navigations, the system provides several functions that assist users in walking around the city. The street picture function allows users to point to any location on the map and see street view pictures near the location. They can use the note function to mark their own points of interests or destinations on the map.

The *Kansei* stroll map system runs on both desktop/laptop and tablet PCs. The authors assumed that the users plan routes before visiting a city using the system on desktop/laptop PCs and the system on the tablet PC helps them to find places having unique impressions when they are walking in the field.

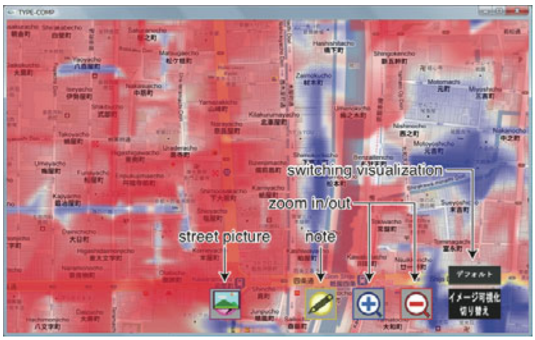


Fig. 4. Screen configuration of the *Kansei* stroll map

4 User Study of the *Kansei* Stroll Map

4.1 Methods

After implementing the *Kansei* stroll map, a user study conducted in Kyoto investigated how the map affects user actions and decisions. Four groups (Groups 1–4), each consisting of two undergraduate students known to each other, participated in the studies. None of the participants had ever lived in Kyoto and had not visited Kyoto in the past three years. The participants used the map before going to Kyoto for planning their walking routes and during their walks in the city.

The studies used two different areas of the city: the area surrounded by Karasuma St., Nijo St., Kawaramachi St. and Shijo St. (area A) and that surrounded by Kawaramachi St., Nijo St., Higashioji St. and Shijo St. (area B). In area A, the participants started walking from Kawaramachi station, at the corner of Kawaramachi and Shijo streets, and walked toward Karasuma Oike station, at the corner of Karasuma and Oike streets. In area B, they walked from the Yasaka shrine at the corner of Higashioji and Shijo streets to the city hall at the corner of Kawaramachi and Oike streets. The participants walked around one area using the *Kansei* stroll map with visualized impressions of the streetscapes and the other area using the map without visualizations. The groups were assigned the combinations of areas and conditions shown in Table 2.

Pre-planning. Before going to Kyoto, the participants were given access to the *Kansei* stroll map on a laptop PC and a commercially available paper-based map. An experimenter gave them an overview of the user studies and instructed them in the operation of the *Kansei* stroll map. Each group used the set of maps to make their approximate walking plans for both the areas. Meanwhile, the *Kansei* stroll map provided visualizations only for one area, according to the conditions shown in Table 2. After making the walking plans for both the areas, the participants answered a questionnaire, consisting of several evaluation criteria with a Likert scale and free form responses.

Walking in Kyoto. While walking in Kyoto, the participants used the *Kansei* stroll map running on a tablet PC and a paper-based map. An experimenter again instructed them in the use of the stroll map and they moved to the starting point. They walked around the area for 90 minutes using the set of maps. After the first walk, they moved to the next area and walked there in the same manner. The *Kansei* stroll map provided visualizations only for one area, as planned. When walking, the participants were not required to walk the area the way they had planned, but were allowed to freely change their route, depending on their interests. An experimenter accompanied each group, recording the group's video and a GPS logged their route. After the walk, the participants answered a questionnaire.

Table 2. Conditions and areas assigned for each group

Group	First walking	Second walking
Group 1	Area A with visualized impression	Area B without visualization
Group 2	Area B with visualized impression	Area A without visualization
Group 3	Area B without visualization	Area A with visualized impression
Group 4	Area A without visualization	Area B with visualized impression

4.2 Results of Pre-planning

In their planning, most participants looked for places to visit using the maps. When planning without visualizations, they tended to select famous places from the map; however, with visualization, they used the visualized impressions on the stroll map to find places or streets that interested them. The participants in Group 1 checked the visualized impressions of the pair *noisy–quiet* on the *Kansei* stroll map, as they wanted to walk through quiet districts and visit small temples in area A. The participants in Group 2 wanted to visit temples and cultural areas; they looked for impressions of areas with the adjectives *traditional* or *fantastic* on the stroll map. They planned to visit the Gion district as well as the Enkoji and the Yoboji temples in area B. Group 3 checked the impressions of *noisy–quiet* on the *Kansei* stroll map because they wanted to go shopping in area A. They found a narrow street marked with dense red colours on the map and decided to walk there. This is Nishiki-koji St. and it contains many small shops. Group 4 selected the names of chain stores to identify their locations as landmarks; however, they did not plan any specific route or place to visit in either area, even though they checked the visualization of the *modern–traditional* impression for area B.

4.3 Results of Walking in Kyoto

Figures 5(a) and 5(b) show the routes of all the groups in areas A and B, respectively. In these figures, solid lines indicate the routes of the groups using visualized impressions while dashed lines indicate the routes of the groups without using the visualizations. The circle and square represent the start and destination points, respectively. The participants in Groups 1–3 walked toward the places or streets they checked during their planning before visiting Kyoto. Occasionally, they found something interesting and changed their route. The participants in Group 4, who did not check any specific route or place beforehand, chose their direction by looking at surrounding areas. In all groups, the participants frequently checked the maps, mostly to determine their position and to determine the direction to their destination. In some cases, they used the visualizations on the *Kansei* stroll map to compare the streetscapes and the visualized impression, but not frequently.

5 Discussions

5.1 Frequency of *Kansei* Stroll Map Usage

The participants used the visualization of the *Kansei* Stroll map more frequently during the pre-planning process than in the field. In the planning process, the information that participants used was limited to maps, pictures and names of tourist attractions. It is difficult to form impressions of the places and streets from this information and the visualization of the impression helped the participants get a sense of the characteristics. In contrast, while walking, the participants directly perceive various kinds of information, such as scenery, confluence, sounds and smells, from their surroundings. They gave priority to this perceived information over that visualized on the map. Thus, the visualization on the *Kansei* stroll map proved most effective for planning the trip.

In the questionnaire, the average score for the question ‘The visualization of the *Kansei* Stroll map was useful for the planning’ was 4.13 (standard deviation 0.64) where 1 represents ‘strongly disagree’ and 5 represents ‘strongly agree.’ The score for the question ‘The visualization of the *Kansei* Stroll map was useful for walking the city’ was 3.50 (standard deviation 0.76). The results suggest that the participants found the visualized impression more useful during the planning process. One participant commented: ‘The stroll map was effective for knowing the atmosphere of places before visiting there.’

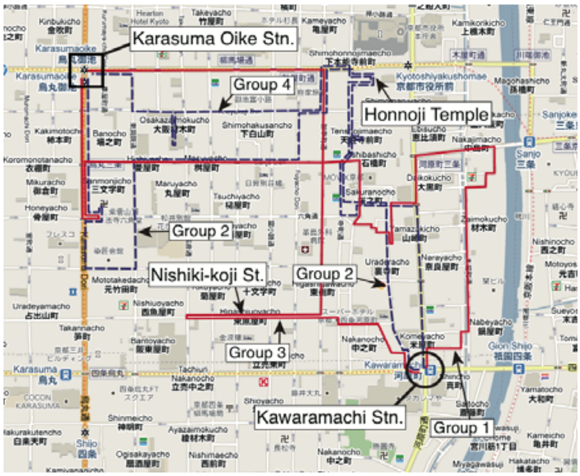
5.2 Route Comparison

In Fig. 5, there are some areas visited by two or more groups and areas visited by only one group. When without using the visualizations, both groups commonly visited the Honnoji temple in area A and walked through the riverside of Kamogawa River in area B. To compare the similarity of the routes between the conditions, the authors defined a route-overlapping rate that is expressed by

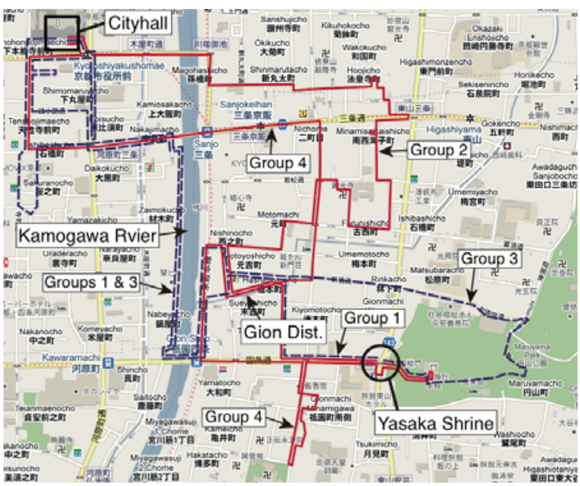
$$overlap_r = \frac{o_r}{d_r} \tag{2}$$

where d_r represents the total distance of the route r while o_r represents the total distance that route r overlaps the routes of other groups.

Figure 6 shows the route-overlapping rate for each group and each condition. The rate is higher when the participants used the maps without visualization than that used with visualization, except for Group 4. Since the names of the Honnoji temple and Kamogawa Riverside are famous, participants tend to focus on them, especially



(a) Area A



(b) Area B

Fig. 5. Walking routes during the studies in Kyoto

without visualization. The participants usually tried to find effective routes that directly connect famous places or tourist attractions. This led the participants to plan similar walking routes when they had no visualization. On the other hand, with the visualized impressions, participants easily found non-famous places or streets, which could be difficult to find with a normal map. They planned to stray from the route based on their interests. These results suggest that the stroll map increases the variety of routes that tourists choose.

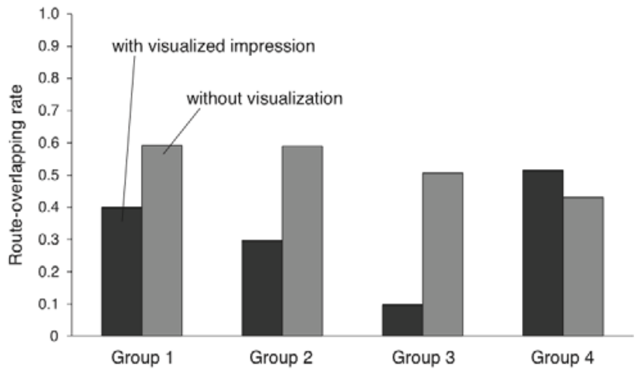


Fig. 6. Route-overlapping rate for each group and each condition

6 Conclusions

This paper introduced a *Kansei* stroll map based on the visualization of impressions of streetscapes. User studies conducted for pre-planning and while walking the city led the authors to the following findings. First, the *Kansei* stroll map helps tourists find places or streets that match their personal interests. Second, the visualized impression is especially useful for pre-planning. Third, the stroll map increases the variety of walking routes that tourists choose. Although the user studies in this paper used only a simple map system with visualized impressions, the results show that the system provides useful cues for walking a city. Future studies will address the integration of impression visualizations and conventional navigation methodologies to support various styles of travel.

Acknowledgments. This study was supported by a grant from The Ministry of Education, Culture, Sports, Science and Technology, Japan (Grant-in-Aid for Scientific Research 20700197).

References

1. Cheverst, K., Davies, N., Mitchell, K., Friday, A., Efstratiou, C.: Developing a Context-aware Electronic Tourist Guide: Some Issues and Experiences. In: Proceedings of ACM CHI 2000 Conference on Human Factors in Computing Systems, pp. 17–24 (2000)

2. Hinze, A., Buchanan, G.: Context-awareness in Mobile Tourist Information Systems: Challenges for User Interaction. In: Proceedings of Workshop on Context in Mobile HCI 2005, 4 pages (2005)
3. Google Maps, `http://maps.google.com`
4. Kinoshita, Y., Nakama, T.: A Proposal of the Kansei Structure Visualization Technique for Product Design. In: Proceedings of the International Conference on Kansei Engineering and Emotion Research 2010, pp. 324–331 (2010)
5. Nakama, T., Kinoshita, Y.: A Kansei Analysis of the Streetscape in Kyoto - An Application of the Kansei Structure Visualization Technique. In: Proceedings of the International Conference on Kansei Engineering and Emotion Research 2010, pp. 314–323 (2010)
6. Nagamachi, M.: Kansei Engineering: A new ergonomic consumer-oriented technology for product development. International Journal of Industrial Ergonomics 15(1), 3–11 (1995)
7. Osgood, C.E., Suci, G.J., Tannenbaum, P.H.: The Measurement of Meaning. University of Illinois Press (1957)

Multivariate Data Visualization: A Review from the Perception Aspect

Yan Liu

Department of Biomedical, Industrial, and Human Factors Engineering
Wright State University, Dayton, OH 45431
yan.liu@wright.edu

Abstract. Last two decades have seen the development of varieties of new methods for visualizing multivariable data and a few attempts have been made to survey and compare some of these techniques. Despite valuable reviews, they did not systematically study the perception tasks involved in these techniques which affect the efficiency of their information decoding due to limited humans' visual perception system. This paper serves to fill this gap through reviewing three well known multivariate visualization techniques apropos their associated perception tasks for three data exploration purposes. Advantages and disadvantages of each tool are discussed.

Keywords: visual perception, data exploration, trellis display, parallel coordinates, pixel-based techniques.

1 Introduction

Visualization is crucial to data analysis; it provides a front line of attack, revealing intricate structure in data that cannot be absorbed in any other way [1]. Fig. 1 shows a simplified conceptual model of the data visualization process. The most crucial step in this process is visual encoding which converts a data table into graphic *G* through mapping the variables of the data table to some visual elements (*VE*). Finally, users have to perform perception tasks to decode the visual presentation in order to extract data characteristics. Because the goal of data visualization is to facilitate users in understanding data and its message rather than the graph itself, the perception tasks involved are key factors to effective visual designs. Visualization mainly serves three types of data analysis tasks: exploration, confirmation, and presentation. This article focuses on the data exploration purpose of visualization.

Last two decades have seen the development of varieties of new methods for visualizing multivariable data and a few attempts have been made to survey some approaches [2,3,4]. Despite valuable reviews, they only presented summaries of the principles of these techniques without systematically studying the involved perception tasks which affect their efficiency of information decoding due to limited capacities of humans' perception system. In addition, works have been done to compare some techniques in terms of their support to certain data exploration tasks using several datasets [5,6]. These comparisons, however, did not provide much insight into the

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 221–230, 2011.

root causes of the pros and cons of these visualization tools. It is the nature of the involved perception tasks that determines the success of exploratory purposes. In this research, a perception task is considered efficient if its intended data exploration task can be accomplished quickly and accurately.

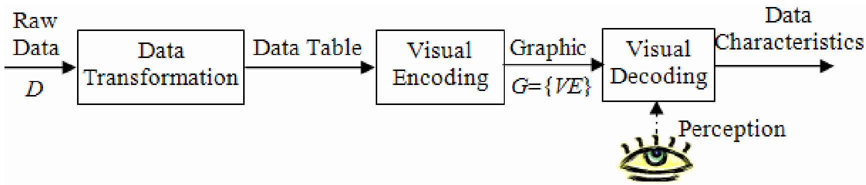


Fig. 1. Conceptual Model of the Data Visualization Process

Within this context, this article aims at reviewing multivariate visualization techniques apropos their associated perception tasks for data exploration purposes. Due to the space constraint of the article, however, only three techniques – trellis display, parallel coordinates, and pixel-based display – are examined here. They are chosen for the study because they are well recognized in the information visualization field, they allow each data point to be plotted in the figure, and they can be used to visualize both continuous and discrete data. In addition, this project focuses on three typical genres of data exploration purposes: identifying the dependence between variables, detecting outliers, and discovering classification rules of a predefined class variable in the data.

2 Description of Datasets

The dataset used to illustrate the discussions in following sections is adapted from the "auto-mpg" dataset retrieved from the UCI machine learning repository [7]. Five attributes from the original auto-mpg dataset are used here, including *mpg*(continuous), *cylinders*(3/4/5/6/8), *horsepower*(continuous), *weight*(continuous), and *origin* ("American", "European", and "Japanese"). *origin* is considered as the class variable when the classification task is discussed. In the following figures shown for this dataset except the pixel-based display, red, blue, and green data points represent the American, European, and Japanese cars, respectively. There are totally 392 data records in the dataset.

3 Visualization Techniques

In each of the three visualization techniques discussed in this section, its mappings of variables, perception tasks involved in the three chosen data exploration purposes, and a discussion of the implications of its merits and perception issues are presented.

3.1 Trellis Display

The basic idea behind Trellis display is to show any one of a large variety of 1-D, 2-D and 3-D statistical data view plot types in an automatically generated trellis layout of panels, where each panel displays the selected plot type for a slice on one or more additional discrete or continuous conditioning variables [8]. Panels are laid out into columns, rows and pages. Fig. 2 illustrates the trellis display for the *auto-mpg* dataset.

Mapping of Variables. The mapping of variable X_i depends on its role in the display as an axis variable, conditioning variable, or superposed variable. If X_i is an axis variable, then it is mapped to one of the coordinates in panels. If X_i is a conditioning variable, then each panel represents one of its levels (for categorical variables) or intervals (for continuous variables). Continuous variables have to be divided into intervals (usually overlapped) in order to be used as conditioning variables in trellis displays. If X_i is a superposed variable, then it is mapped to colors or symbols of points in panels. For example, in Fig. 2, *mpg* and *horsepower* are axis variables, *cylinders* and *weight* are conditioning variables, and *origin* is a superposed variable. As *weight* is a continuous variable, it is shingled into 5 intervals with 1/10 overlapping. *weightE* is used to represent the shingled variable. Each row in Fig. 2 represents a value of *WeightE*. The values of *cylinders* and *weightE* increase from left to right in columns and from bottom to top in rows, respectively.

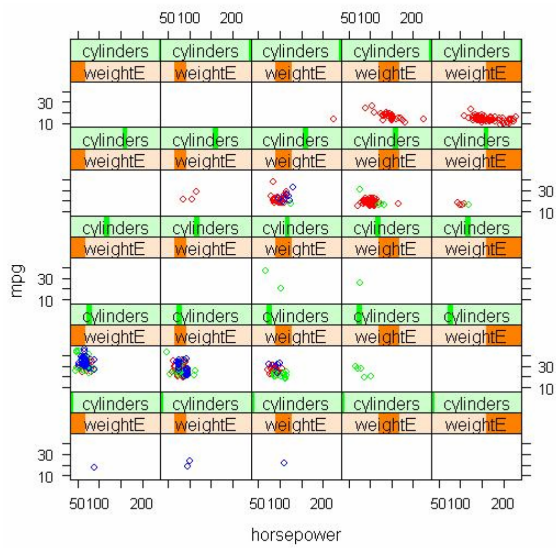


Fig. 2. Trellis display of the auto-mpg dataset

Perception tasks involved in data exploration

Dependency between Xi and Xj. If X_i and X_j are axis variables and X_c is the set of all conditioning variables, then the perception tasks involved in discovering the dependency between X_i and X_j conditioning on X_c are the same as those in a scatterplot.

To examine the dependency between X_i and X_j conditioning on a subset of X_c requires mentally merging panels associated with the same conditioning variables under study into an integrated scatterplot and then examining the dependency between X_i and X_j in the integrated plot. For instance, to identify the dependence between *mpg* and *horsepower* among cars with 8 cylinders in Fig. 2, one needs to mentally integrate panels at the right-most column into a scatterplot and then examine the dependency between *mpg* and *horsepower* in the integrated plot. It is not difficult to perceive from Fig. 2 a negative correlation between *mpg* and *horsepower* for cars with 8 cylinders. However, more detailed information on their dependency, such as the linearity, is difficult to perceive due to the panel merging process.

Outliers. The perception task for detecting outliers in the combination of all conditioning variables and one or both of axis variables is to spot points whose positions do not conform to the overall patterns of the dependency between the axis variables in each panel. For example, in the second right-most panel at the second-top row in Fig. 2, an outlier in the combination of *mpg*, *weight*, and *cylinders,* which has a relatively higher value of *mpg*, can be easily identified.

To detect outliers in the combination of some but not all conditioning variables and one or both axis variables, one needs to first mentally merge panels associated with the same combination of the conditioning variables of interest into one scatterplot and then spot the outlier in the integrated plot.

The perception task for detecting outliers in the combination of all conditioning variables without considering axis variables is to check whether some panels only have few data points. However, if some conditioning variables are shingled continuous variables, then the number of points in panels is greatly affected by the shingling parameters of these variables.

Classification. The perception task for classification based on the combination of all the conditioning variables (with or without axis variables) is to compare the densities and/or areas covered by data points with different colors in each panel. For example, Fig. 2 shows that all cars that have 8 cylinders and are the heaviest (the right-most panel at the top row) are American cars.

However, for classification based on an individual variable or a subset of all conditioning variables, one needs to first mentally integrate the associated panels for points at the same interval or level of the individual variable or the same combination of the variables of interests and then compare the densities of and/or areas covered by points with different colors in the integrated panel.

Implications

As seen from the above discussions, the efficiency of perception tasks involved in trellis displays are greatly affected by the assignment of variables. The major advantage of the trellis display, from the perception perspective, seems to be its effective demonstration of the information on the axis variables taking into consideration of all the conditioning variables because the associated perception tasks would become the same as those in our familiar scatterplots. In other cases, however, the process of mentally integrating panels associated with the variables of interest is necessary, which can make the corresponding perception tasks very inefficient. Besides, due to the trellis display's shingling mechanism which categorizes continuous variables into several overlapping intervals, some information regarding the original continuous variables may not be demonstrated adequately, such as the linearity of the dependency between them and other variables and the their outliers.

3.2 Parallel Coordinates

Parallel coordinates, pioneered by Inselberg [9], is a well-known technique for visualizing multivariate data. In this approach, each variable is represented by a vertical axis and the m variables are organized as uniformly spaced vertical lines. A data record in m-dimensional space manifests itself as a connected set of points, one on each axis. Fig. 3 shows the parallel coordinates of the *auto-mpg* dataset.

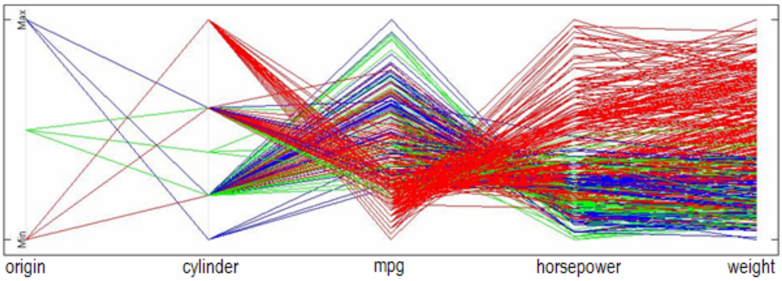


Fig. 3. Parallel coordinates of auto-mpg dataset

Mapping of variables

Variable X_i is represented as the ith vertical axis in a 2D space and its values are usually scaled so that the maximum and minimum values correspond to the top and bottom points on its axis, respectively. Data record $R_i = \{X_1 = r_{i1}, X_2 = r_{i2}, \ldots, X_m = r_{im}\}$ is represented as a set of m-1 connected line segments which connect to vertical lines at the corresponding variables' values.

Perception tasks involved in data exploration

Dependency between Xi and Xj. If the axes of X_i and X_j are adjacent in the plot and they both are continuous, then to find their dependency is to compare the directions of line segments between their axes. If most lines are parallel, they are positively correlated. On the other hand, if most lines cross over one another, they are negatively correlated. However, if the lines form a figure with a hyperbolic envelop, then these two variables are not correlated. For instance, in Fig. 3, crossings of line segments between the *mpg* and *horsepower* axes whose intersection points are not close to the center of the space between their axes suggest a negatively nonlinear correlation between them. Besides, most line segments between the *horsepower* and *weight* axes seem to be parallel, which indicates a positively linear relationship between them. If X_i is discrete and X_j is continuous, then identifying their dependency is to compare the lines which connect to different regions on the axis of X_i in terms of their distributions on the axis of X_j.

If the axes of X_i and X_j are nonadjacent in the plot, on the other hand, then the viewer has to first locate the corresponding points on the axes X_i and X_j by tracing line segments before examining the dependency between X_i and X_j. This line tracing process makes the perception task of identifying the dependency between X_i and X_j, very inefficient because the information is processed in the manner of controlled processing (as opposed to preattentive processing). Besides, it becomes almost impossible to trace corresponding points on nonadjacent axes when many connection lines or points overlap.

Outliers. To detect outliers in individual variables is to check whether there are few connection points on their axes that are well separated from others.

For detecting outliers in the combination of two variables whose axes are adjacent in the plot, one needs to examine whether there are line segments between them whose directions do not agree with the overall patterns of other line segments. To detect outliers in the combination of two or more variables whose axes are nonadjacent in the plot, one has to first trace line connections to find corresponding points on their axes and then check whether, among connection points located at the at some regions on the axes of these variable, there are few points that are far away from others. The process of line tracing makes the perception task very inefficient.

Classification. Discovering classification rules based on one variable is to compare the densities of lines with distinctive colors that connect to the same regions on its axis. For example, to classify the *origin* of cars based on *cylinders* in Fig. 3, one needs to compare the densities of red, blue, and green lines connected to each of the five points on the *cylinders* axis. Examining the line segments between the *cylinder* and *mpg* axes, one can see that all cars with 8 cylinders and most cars with 6 cylinders are American cars and all cars with 5 cylinders and 3 cylinders are European cars and Japanese cars, respectively.

Discovering classification rules based on the combination of two or more variables requires first tracing line connections to find lines whose connection points on the axes of these variables satisfy the same combination of these variables and then comparing the densities of lines with distinctive colors among them. If two variables are adjacent, then the directions of the line segments between them can help locating corresponding connection points.

Implications

The ease of implementation and interpretation makes parallel coordinates a focus of attention since its inception. From the above discussions we can see that in most data exploration tasks, the effectiveness of parallel coordinates is significantly affected by the severity of the overlapping of line segments and the order of the axes of variables.

By taking advantages of humans' capacities in comparing line directions, preattention to orientations, and Gestalt proximity theory, parallel coordinates allow one to effectively perceive the dependency between two continuous variables whose axes are adjacent and outliers in individual variables and in the combination of two variables whose axes are adjacent. However, because the points associated with the same data record on different axes are identified through connected line segments, the order of axes can greatly impact the efficiency of the perception tasks engaged in data exploration undertakings in which locating corresponding points on different axes is required.

3.3 Pixel-Based Visualization

In pixel-based visualization techniques, each data point is represented with a colored pixel. Several pixel-based visualization approaches have been proposed depending on the purpose of the visualization, how the pixels are arranged and the shape of the display window [10,11]. This paper, however, only discusses the spiral pixel-based visualization technique and the perception tasks involved. Figure 4 shows the spiral pixel-based display of the auto-mpg dataset. The top-left window is to visualize the relevance factor, the inverse of the combined overall distances, of data points in all the five variables of the dataset.

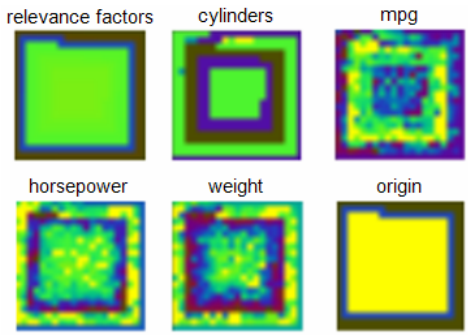


Fig. 4. Spiral pixel-based display of the auto-mpg dataset

Mapping of variables and data records

Variable X_i is represented as the *i*th square window in a 2D space which is filled with colored pixels starting from its centre. Data record $R_i = \{X_1 = r_{i1}, X_2 = r_{i2}, \ldots, X_m = r_{im}\}$ is represented as a set of *m* pixels, each in one window associated with a variable; the relative locations of these pixels are the same in all windows. The color of a pixel demonstrates its corresponding data value. The colors map, ranging from yellow over green, blue, and red to almost dark, indicates increasing differences between the data values that pixels represent and the user's specified queries because of the seemingly lessening in the lightness of these colors with the same saturation [10].

Perception tasks involved in exploratory tasks

Dependency between Xi and Xj. If both X_i and X_j are continuous, then to identify their dependency is to compare the colors of the corresponding pixels in their windows. If the color patterns in the two windows are similar according to the color map, then they are positively correlated. If the color patterns are opposite in the two windows, (i.e., light colors in one window correspond to dark colors in the other window), then they are negatively correlated. For example, Fig. 4 shows that the light pixels in the *mpg* window correspond to dark pixels in the *horsepower* windows, which suggests a negative correlation between *mpg* and *horsepower*. However, because each color in the color map represents an interval of values rather than a single value, the linearity of the dependency between X_i and X_j cannot be perceived adequately from the pixel-based display.

If X_i is discrete and X_j is continuous, then to identify their dependency is to compare the distributions of pixels in X_j's window whose corresponding pixels are of different colors in X_i's window. For instance, to identify how the distribution of *mpg* depends on the level of *cylinders*, one needs to compare the distributions of pixels in the *mpg* window whose corresponding pixels are yellow, green, blue, violet, and dark-olive in the *cylinders* window. Fig. 4 shows that for yellow and green pixels in the *cylinders* window (mainly at the center or edge of the window), their corresponding pixels in the *weight* window are mostly yellow and green. This suggests all cars with 3 or 4 cylinders have relatively low values of weight.

Outliers. To detect outliers in individual variables is to spot few pixels whose colors are different from others in their corresponding windows. To detect outliers in the combination of two or more variables is to discover pixels whose colors do not agree with the general color patterns in the windows of these variables. One needs to compare the colors of the corresponding pixels in these windows in order to find the pattern, but matching pixels at the same relative locations in different windows is an inefficient operation, especially when the windows are far apart from each other.

Classification. To discover classification rules based on one variable or the combination of two or more variables is to compare the areas of the pixels which have the same colors in its or their window(s) but correspond to different classes. The class labels of pixels are identified through the colors of the corresponding pixels in the window of the class attribute. For example, in Fig. 4, we can see from the *origin* window that yellow, blue, and dark-olive regions represent American, European, and Japanese cars, respectively. In order to classify the *origin* of cars based on *cylinders*, one needs to, for the yellow, green, blue, violet, and dark-olive pixels in the *cylinders* window, respectively, compare the areas of pixels that correspond to different *origin* classes. As shown in the *cylinders* window, all its yellow pixels, representing 3-cylinder cars, correspond to the Japanese cars (the dark-olive pixels in the *origin* window), and all the dark-olive pixels in the *cylinders* window, representing 8-cylinder cars, correspond to American cars (the yellow pixels in the *origin* window). However, the green pixels in the *cylinders* window, representing 4-cylinder cars, correspond to pixels with all three colors in the *origin* window, which means that simply knowing a car has 4 cylinders cannot classify its origin adequately. However, it is very inefficient to classify the *origin* of cars based on *mpg* or the combination of *mpg* and *weight* because pixels with the same colors are scattered in the *mpg* and *weight* windows.

Implications

Using one pixel to represent a data point in the pixel-based technique allows displaying the maximum possible number of points in one screen without overlapping of points.

Colors of pixels demonstrate the values of their corresponding data points in the pixel-based technique. However, because colors do not provide quantitative information effectively, viewers may have to always refer to the color map for the actual values that the colored pixels correspond to, especially for continuous variables. Besides, in the pixel-based display, continuous variables are literally discretized, one color for an interval of values; this prevents perceiving the linearity of the dependency between two continuous variables. How the intervals are divided also affects the discovery of clusters.

In the pixel-based display, relating corresponding pixels in different windows requires the viewer to mentally match their relative positions, both horizontally and vertically, in these windows; this task can very inefficient when the distributions of the pixels do not demonstrate clear structures and/or these windows are far apart and/or not aligned horizontally or vertically.

4 Conclusions and Future Research

It is commonly believed that visualization tools can help unveiling hidden patterns and relationships among variables. Recently, there has been growing interest in developing tools for visualizing multivariate data in response to the needs of handling complex datasets in various fields. So far, most practices in evaluating visualization techniques involve only administering them on several datasets and examining the demonstrated patterns. Such an approach, however, does not provide much insight into the underlying perception issues which determines their effectiveness in revealing information to the end users. No matter how clever the choice of the information, and no matter how technically impressive the encoding, a visualization fails if the decoding fails [1].

This paper has examined three well recognized multivariate visualization techniques with respect to their associated perception tasks for three data exploration purposes. Some advantages and disadvantages of each technique are discussed.

My future work will involve conducting more rigorous formal experimentations to test the findings from this study. Examining the perception tasks involved in visualization techniques, both theoretically and experimentally, gives a scientific foundation on the evaluation of their effectiveness. Later on, research will be conducted on comparing the effectiveness of these techniques to help developing guidance in choosing appropriate tools based on the characteristics of datasets and exploration tasks.

References

1. Cleveland, W.S.: Visualizing Data. Hobart Press, New Jersey (1993)
2. Wong, P.C., Bergeron, D.: 30 Years of Multidimensional Multivariate Visualization. In: Nielson, G., Hagan, H., Muller, H. (eds.) Scientific Visualization – Overview, Methodologies, and Techniques, pp. 3–33. IEEE Computer Society Press, Los Alamitos (1997)
3. Tatu, A., Albuquerque, G., Eisemann, M., Schneidewind, J., Theisel, H., Magnork, M., Keim, D.: Combining Automated Analysis and Visualization Techniques for Effective Exploration of High-Dimensional Data. In: 2009 IEEE Symposium on Visual Analytics Science and Technology, pp. 59–66 (2009)
4. Keim, D.A.: Information Visualization and Visual Data Mining. IEEE Transactions on Visualization and Computer Graphics 8(1), 1–8 (2002)
5. Keim, D.A., Kriegel, H.-P.: Visualization Techniques for Mining Large Databases: A Comparison. IEEE Transactions on Knowledge and Data Engineering, Special Issue on Data Mining 8(6), 1–29 (1996)
6. Grinstein, G., Patrick, E., Pickett, R.M., Laskowski, S.J.: Benchmark Development for the Evaluation of Visualization for Data Mining. In: UsFayyad, U., Grinstein, G., Wierse, A. (eds.) Information Visualization in Data Mining and Knowledge Discovery, pp. 129–176. Morgan Kaufmann Publishers, San Francisco (2002)
7. UCI Machine Learning Repository Irvine, CA: University of California, School of Information and Computer Science, `http://archive.ics.uci.edu/ml`

8. Cleveland, W.S., McGill, R.: Graphical Perception: The Visual Decoding of Quantitative Information on Statistical Graphs (with Discussion). Journal of the Royal Statistical Society Series A 150(3), 192–229 (1987)
9. Inselberg, A.: The Plane with Parallel Coordinates. The Visual Computer 1(1), 69–97 (1985)
10. Keim, D.A.: Designing Pixel-Oriented Visualization Techniques: Theory and Applications. IEEE Transactions on Visualization and Computer Graphics 6(1), 59–78 (2002)
11. Hao, M.C., Dayal, U., Keim, D., Schreck, T.A.: Visual Analysis of Multi-Attribute Data Using Pixel Matrix Displays. In: Proceedings of the SPIE - The International Society for Optical Engineering Visualization and Data Analysis, vol. 6495, p. 649505-1-9 (2007)

Methods for Service Sciences from Visualization Points

Hiroyuki Miki[1], Naotsune Hosono[2], and Sakae Yamamoto[3]

[1] Oki Electric Ind. Co., Ltd., Software Center
1-16-8 Chuou, Warabi-shi, Saitama 335-8510, Japan
[2] Oki Consulting Solutions Co., Ltd.
[3] Tokyo Univ. of Science
hmiki@cf.netyou.jp, hosono903@oki.com, sakae@ms.kagu.tus.ac.jp

Abstract. Recently, situations surrounding business organizations have been changing rapidly and extensively. Due to the wide availability of broadband Internet in many countries, international division of labor, namely offshoring, has become quite common, and global competition is growing intense [1]. As a result, business organizations must seek to produce added-values and higher efficiency to survive the competition. To produce added-values and higher efficiency, a variety of issues have been coped with in the research area named Service Sciences or Service Engineering (SS in short). Since a large amount of data is acquired, analyzed, and utilized in SS, a scientific and engineering approach is important in it. This paper firstly outlines the approach and visualizations for it, and then discusses good solutions for both the service provider and the customer. By addressing “visualization”, important points of SS activities are clarified.

Keywords: Service Sciences, Customer Expectation Management, Persona, Scientific Disciplines, Productivity.

1 Introduction

With development of economy, the ratio of service-based industries against other industries such as manufacturing industries rises. These days, this change is more evident by industry’s aggressive business efforts such as outsourcing for more business productivity [1].

However, achieving higher service productivity is not easy. While manufactured products have visible and tangible shapes, services doe not have such properties. They are not all visible. Their production and consumption happen at the same time and disappear instantly. They vary with situations, and so on. Hence, they are more difficult to measure and visualize than manufactured products, which poses difficulties of grasping service process and tends to result in lower productivity.

SS is a research approach which applies various scientific and engineering methods to tacit knowledge and tactics of services and aims at higher service productivity by process improvement, cost cut, and so on along with higher satisfaction [2, 3, 4, 5]. SS approach is characterized as the following.

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 231–238, 2011.

- Massive data handling by scientific and engineering methods: In SS, Scientific and engineering methods are used to process massive data with reference to innovation theories [6, 7] and other theories in Business Administration, Psychology, and other human science areas.
- Aiming at productivity improvement: Previously improvement of productivity highly relied on worker's improvement of ways of work, and was hence implicit or unsystematized. In SS, service productivity is improved based on the notion that added-value divided by labor input is service productivity.
- Coping with services, customer's benefit, and customer's expectation: Service means an activity which gives some benefit to the customer. Since customer's benefits arise in customer activities, analyses of such activities, in other word human-centered viewpoints, are important. To bring customer some benefit, management of customer's expectation is important.

In the following, in section two, purpose of visualizations in SS is explained with reference to the typical scientific process named Optimal Design Loop. Then visualizations related to each of the above three characteristics are explained: basic visualizations of massive data in section three, visualizations of service productivity in section four, and visualizations of customer's benefit (expectation) in section five. In section six, good solutions for both the service provider and the customer are discussed.

Note that the term "visualization" is used in this research as a general term of making things visible. It includes descriptions of field research results in the social science area as well in addition to data visualization in information science and engineering areas.

2 Optimal Design Loop and Purpose of Visualizations

As a scientific and engineering method, Optimal Design Loop is gaining acceptance for its efficiency and effectiveness in the field of SS (Fig. 1) [3]. This design loop consists of processes with data observation and accumulation, analysis, design, and application. The advantage of the loop is that once the prediction model is built using the data relation analyzed from observation and accumulation, prediction of the entire data will be possible even if it is later given only partial observation data. By repeating the loop in such a process, precision enhancement and expanded prediction area are eventually expected.

In Optimal Design Loop, visualization of data is important along with observation and accumulation of data. One reason of this is that when analyzing regularity and relations of a large amount of observed and accumulated data, it will be very hard to find such characteristics without visualization. With visualization, different viewpoints can be seen easily.

There is also a close relationship between visualization of data and observation and accumulation of data regarding qualitative and quantitative data. If observed and accumulated data is qualitative, its analyzed results will be qualitative unless it is converted to quantitative data. The same is true in the opposite case. Such conversion is important for customer's better understanding of the data. In the body fat monitor, raw qualitative data are converted to another qualitative data based on the prediction model.

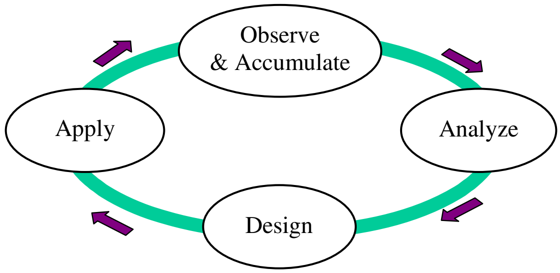


Fig. 1. Optimal Design Loop in Service Sciences. (This is created based on Motomura [3]. The emphasized steps will be different depending on the types of objects and innovations.)

It is also important what the visualization is for and whom it is for. Although Optimal Design Loop is mainly for the creation of service, it must be visualized for customers' easier understandings. If the visualization is useful for the customer, it will motivate the customer to continue to use the service and product ("service" in short hereafter) and will lead to continual loop of the optimal design loop. The continual loop will bring the service provider a better chance to improve the service and bring the customer a better chance to get an improved service.

3 Visualizations Related to Massive Data Handling

An example of SS is a scale with a body fat monitor found on the market today. In "Observe & Accumulate" phase of Fig. 1, information about body weights, electric conductivities and body fat percentages from hundreds of people are observed and collected in a database. Then in "Analyze" phase, the analyzed data formulate the relations of body information. In "Design" Phase, the manufacture designs a scale with a body fat monitor using the formula. In "Apply" phase, the formula is applied when the scale is used to calculate body fat percentages, the skeletal muscle rate, and the basal metabolism based on the weight and the electric conductivities. At the same time, the observation data uploads to accumulate by means of the Internet in "Observe & Accumulate" phase. Then the analyzed data contribute to the next generation service with enhanced precision and expanded functions.

When accumulation and analysis of quantitative mass data clarify actual human activities, and prediction precision is improved by rotating the cyclic loop, creation of new added-values and efficiency improvement can be expected.

In each phase, various visualizations are used to understand and utilize tendencies of information for both the service provider and the customer.

4 Visualizations Related to Service Productivity

Service productivity which SS aims at is approximately defined by the following formula. Service productivity is added-value divided by labor input [3].

$$\text{Service productivity} = \text{Added-value} / \text{Labor input} \quad (1)$$

Following the formula in (1), the improvement of service productivity requires improvement in the denominator, the numerator or both conditions. SS is expected as a way to bring significant progress by improving both the denominator and the numerator. Namely, service productivity will rise through less labor input by way of process improvement and cost cutting as well as higher added-values such as higher customer satisfaction.

Since formula (1) is an approximate formula and each element of it can be different measurement and scale, effective visualization is requisite for easier understanding of the relations [4]. Once such relations are explicit, improvement of service productivity can be pursued more easily. In other word, SS is a research effort that seeks for an explicit form of formula (1) and pursues improvement of service productivity based on it. By addressing "visualization", important points of SS activities are clarified.

There are various visualizations related to Service Productivity. For example, for the service provider, they include less labor input to provide a service, more services per hour, higher customer satisfaction with the same labor input, and invested money versus cost cut. On the other hand, for the customer, they include less labor input to receive a service, higher satisfaction with the same labor input, and so on.

As an example, energy consumption monitoring system for IT equipment energy saving, is shown in Fig. 2, 3, and 4 [8]. It is basically installed on an energy administrator's PC and on a user PC, and runs all the time. It monitors energy consumption of user's PC such as power consumption and printed pages, and shows them to the user in a user-friendly manner (Fig. 2). It also judges a total energy saved level based on the monitored data and shows the level as a colored clover ICON, top-left of Fig. 2, for the user's quick understanding of the whole. Energy save ranking can be checked on the website by the user and the energy administrator (Fig. 3).

Fig. 2. Summarized Energy Consumption Data for User [9]

Fig. 3. Website to Show Energy Save Ranking for User and Energy Administrator [9]

Fig. 4. Chronological Power Consumption Data for Energy Administrator [9]

About visualization, firstly different viewpoints are provided: Fig. 2 for the user, Fig. 3 for the user and the energy administrator, and Fig. 4 for the energy administrator. Secondly, quantitative energy consumption data are visualized in Fig. 2 and 5 from raw quantitative data accompanied by some qualitative representations such as the colored clover ICON in Fig. 2. Thirdly, the user is motivated by summarized snapshot data of Fig. 2 and the energy saving ranking in Fig. 3. The energy administrator and the user are related with each other thru different visualizations of massive data and such differences are important for the service, namely energy saving. Compared with energy saving without the system, energy can be saved more with smaller amount of labor input. In this sense, service productivity is higher with the system.

5 Visualizations Related to Customer's Benefit (Expectation)

As stated in section one, "Service" is an activity which gives some benefit to the customer. And such a benefit is one instance of added-value of formula (1) in section four. Thus, it is fundamentally important for SS.

Since benefit is a subjective matter and varies according to the customer, it is very important to manage customer's expectation with which the customer finds a benefit from a given service. If provided service meets customer's expectation, the customer will find some benefit from the service. Otherwise, since the customer will not find a benefit from the service, it can not be a service for the customer in the first place.

To meet customer's expectation, Persona which represents a detailed story and a profile of a group of targeted customers is often used for the design of service (Fig. 5). Thanks to detailed descriptions of Persona, service designer can imagine what is good or not for the Persona and can design details of the service properly [9].

Although visualization plays an important role for this purpose, as the example of energy-saving showed, providing consistent visualizations which meet customer's context-of-use well are not easy. Thus, when creating Persona, it is recommended that before use, in-use, and after use are all clarified as a consistent story and are based on massive data. It is especially so when base technologies for the service are immature and there are big gaps between the story and the supporting technologies. In such a case, Persona's story should be modified again and again until the gaps are minimized enough. After these efforts, balances among service productivity, added-value, labor input, and so on should be considered followed by concentration on productivity.

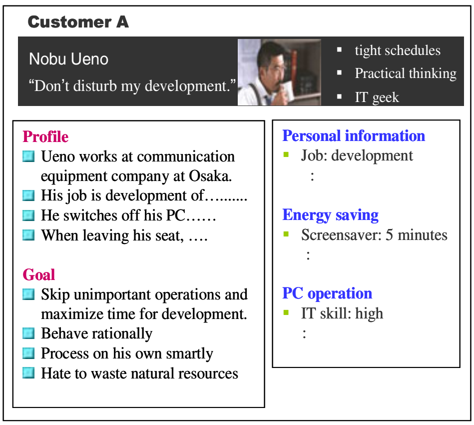


Fig. 5. Example of Persona

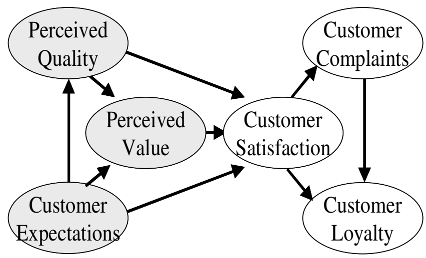


Fig. 6. Customer Satisfaction Index (Arrows represent influence. This is created based on http://www.theacsi.org/index.php)

Once service is established, some evaluation method based on questionnaire such as CSI (Customer Satisfaction Index) (Fig. 6) can be used to evaluate service before, during, and after the use [10]. For example, in CSI, about 20 questions are categorized into six groups: customer's expectations before use, perceived quality, perceived value, and customer satisfaction during use, and customer complaints and customer royalty after use.

Persona and evaluation by questionnaire are both visualized for better understanding of what they should show.

6 Discussion: Toward Good Solutions for Both the Service Provider and the Customer

Most desirable outcome of SS activities should be that both the service provider and the customer are happy thanks to good solutions for both the service provider and the customer. For this, before use, in-use, and after use for each party should be clarified and addressed properly as marketing theories [11] and the notion of User Experience [9] in usability area have pointed out.

However, it is not easy to distinguish what are about the service provider from what are about the customer. Since both added-vale and labor input can be either the service provider or the customer in the formula (1), confusion of the service provider

		Added-value	
		Provider	Customer
Labor input	Provider	Provider issues	Provider & Customer issues
	Customer	Provider & Customer issues	Customer issues

Fig. 7. Productivity of the service provider and the customer

If Expec.	Then Expec.	Then Service
Too high (Low satisfaction)	↘ Less promo. etc.	↗ Service enhancement
Adequate		
Too Low (High or Low satisfaction)	↗ More promo. etc.	↘ Service selection

Fig. 8. Customer expectation management and services

and the customer will easily occur. You may think added-value as value for the customer and labor input as service provider's labor input. However, you may also think added-value as value for the service provider and labor input as customer's labor input, and so on. There are four combinations as shown in Fig. 7.

There is another thing to care about. When a service does not meet customer's expectation, there are several ways to fix it. For example, as shown in Fig. 8, you can change either promotion or service. When customer's expectation is too high and the customer is dissatisfied, you will be required to reduce promotion to lower customer's expectation or you will be required to enhance service, among others.

Since visualizations described in this paper constitute basis to handle these issues, appropriate visualizations are important for the success of service.

7 Concluding Remarks

This paper firstly outlines the approach and visualizations for it, and then discusses good solutions for both the service provider and the customer. By addressing "visualization", important points of SS activities are clarified. Further efforts are expected in this research area.

References

1. Friedman, T.L.: The World Is Flat: A brief history of the twenty-first century. Farrar Straus & Giroux (2005)
2. Maglio, P.P., Srinivasan, S., Kreuren, J.T., Spohrer, J.: Service Systems, Service Scientists, SSME, and Innovation. CACM 49(7), 81–85 (2006)

3. Motomura, Y., Nishida, Y., Mochimaru, M., Akamatsu, M., Naito, K., Hashida, K.: Large Scale Data Observation, Modeling, Service Design and Applying Loop for Service Innovation. Journal of the Japanese Society for Artificial Intelligence 23(6), 736–742 (2008) (in Japanese)
4. Miki, H., Hosono, N., Yamamoto, S.: Researches of Visualization in Service Sciences. SICE 48(5), 391–398 (2009) (in Japanese)
5. Miki, H., Hosono, N., Yamamoto, S.: Transcending human-centered design by service sciences. In: Smith, M.J., Salvendy, G. (eds.) HCI International 2009. LNCS, vol. 5617, pp. 685–692. Springer, Heidelberg (2009)
6. Moore, G.A.: Dealing with Darwin: How great companies innovate at every phase of their evolution. Portfolio (2006)
7. Christensen, C.M.: The Innovator's Dilemma: The revolutionary book that will change the way you do business. Harvard Business School (1997)
8. Someya, T., Murase, K., Tachibana, S.: Cool Clover: Energy consumption monitoring system for IT equipment (2009) (in Japanese) Retrieved from, `http://www.oki.com/jp/otr/2009/n214/pdf/214_r04.pdf`
9. Goodwin, K.: Designing for the Digital Age. Wiley, Chichester (2009)
10. ACSI: American Customer Satisfaction Index Retrieved from, `http://www.theacsi.org/index.php`
11. Kotler, P., Keller, K.: Marketing Management, 13th edn. Prentice-Hall, Englewood Cliffs (2008)

Interacting with Semantics: A User-Centered Visualization Adaptation Based on Semantics Data

Kawa Nazemi, Matthias Breyer, Jeanette Forster, Dirk Burkhardt, and Arjan Kuijper

Fraunhofer Institute for Computer Graphics Research, Fraunhoferstr. 5
64283 Darmstadt, Germany
{Kawa.Nazemi,Matthias.Breyer,Jeanette.Forster,
Dirk.Burkhardt,Arjan.Kuijper}@igd.fraunhofer.de

Abstract. Semantically annotated data gain more and more importance in future information acquiring processes. Especially the Linked Open Data (LOD) format has already experienced a great growth. However, the user-interfaces of web-applications mostly do not reflect the added value of semantics data. The following paper describes a new approach of user-centered data-adaptive semantics visualization, which makes use of the advantages of semantics data combined with an adaptive composition of information visualization techniques. It starts with a related work section, where existing LOD systems and information visualization techniques are described. After that, the new approach will bridge the gap between semantically annotated data (LOD) and information visualization and introduces a visualization system that adapts the composition of visualizations based on the underlying data structure. A case study of an example case will conclude this paper.

Keywords: linked open data, semantic visualization, semantic web.

1 Introduction

Semantic-Web and semantically annotated knowledge and information gain more and more importance in future information and knowledge acquiring processes. While formal descriptions of information, e.g. Ontology or Topic Map are still under the investigation of research institutions and enterprises, semantic structures based on the collective intelligence of web-users became silently an inherent part of the web. Especially the Linked Open Data (LOD) format has experienced a great growth in the open internet and became an established data model for conceptualizing knowledge entities and describing semantic relationships between knowledge entities and domains. The Linked Data format is not only used to model a specific domain by a small set of knowledge engineers, it is more a reflection of the knowledge interpretation of a whole community, which models domain-comprehensive knowledge for structure and disseminate it to a diversified audience. A single Linked Data database gains millions of knowledge entities per day and grows faster than expected.

Although the data in the LOD databases are semantically well-defined, the amount of data is more than sufficient and their structure provides the opportunity for the

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 239–248, 2011.

usage of alternative knowledge-acquisition and interaction with semantics, today's user interfaces of Linked Open Data do not really evince an added value to existing search-result user interfaces. The presentation of the information is often categorized in a concept hierarchy, but the presentation of information is principally text-based.

The already existing trend of alternative interaction ways with computer and information systems is not considered sufficiently in the presentation process of the underlying information. To provide a more natural interaction with the semantic knowledge in the structured LOD databases a graphical representation of the knowledge is valuable. Graphical objects are more similar to the objects of the real world and therefore provide a more intuitive interaction with the semantics structures.

Existing semantics visualization techniques do not consider the surpluses of the Linked Open Data structures, where the semantics structure has to be built-up with a routine of query requests. They focus on various but specific ontology characteristics, e.g. displaying the hierarchical inheritance structure, multiple inheritance or semantic relations between ontology entities. The complex structure of the Linked Data varies, based on the users' query on the data. The heterogeneity of the requested data should be exploited for the visualization and hence enable a more efficient interaction with the underlying semantics.

In this paper we describe a new approach of user-centred data-adaptive semantics visualization for Linked Open Data. The main innovation of our approach is the autonomous selection of different and various visualization types for the results of users' queries. Therefore the results of a user query are analyzed based on different criterions of the data characteristics, e.g. amount of result entities, amount of related entities, amount of related classes and domains or hierarchical structure of the underlying domains. We will show that each result of a semantic user query requires different types of visualization. Complementing our previous work the composition of different visualization types for the visualization of Linked Open Data will be introduced. A classification of the semantic search-result-types will open new possibilities for visualizing and interacting with semantics data. Based on the introduced classification different visualization types and their combination and composition respectively will be introduced. A main aspect of our approach is the interaction with graphical representation of the semantic entities and their structure, which supports the explorative information acquisition process of users and the interaction with alternative devices, especially gesture-based interaction systems.

The paper begins with a related work section, where existing Linked Open Data systems, the amount of their data and their interaction possibilities are analyzed. Supplementary existing semantics visualization techniques are introduced and the interaction possibilities and information acquisition processes are analyzed. In the following section a systematic classification of possible semantic search results will be proceed, where the different characteristics of Linked Open Data are classified especially for their visualization. In comparison to the classification of the semantic attributes the characteristics of information visualization techniques will be introduced and classified. The main section of this paper will bring the Linked Open Data and the visualization characteristics together and introduces a visualization system that adapts the visualization and user interface respectively based on the data under the consideration of users' information acquisition abilities. The paper concludes with an exemplary case study, where different semantic search queries and their visualization will argue the surpluses of our approach.

2 Related Work

The fast growth of the Linked Open Data community has led to a massive amount of interlinked, semantic annotated data on the web, which is called the Web of Data [1]. According to the ‘Linked Data Principles’ by Berners-Lee [2] a lot of open datasets developed and joined the LOD-Community. To use this available massive amount of semantic data, new powerful ways are needed to visualize the semantic information included in the Web of Data and to access the full potential of this new, machine readable data sources. To address this issue, we first give an overview of existing LOD-Systems and in the following introduce two existing semantics visualization systems and their capability to deal with the semantic information in LOD-Systems.

2.1 LOD Databases

This sections describes and analyze two of the most important community based LOD-Databases in the Linking Open Data Cloud[1]. First we will introduce DBpedia[2], a database which was generated using the open data from the web encyclopedia Wikipedia[3] and evolves into a central hub in the LOD-Cloud. In the following we will describe and analyze Freebase[4], a collaborative open database created as a wiki of structured data. To present an overview of existing LOD-Systems, we conclude with a table of common open Datasets, their amount of data and interlinking capability.

DBpedia. DBpedia is one of the most important and fastest growing LOD-Systems in the web of data. It is a community project to extract data from Wikipedia, structure it and make it accessible and interlinked with other datasets. Its great variety of topics led to intersects with many of the available open Datasets, so that it evolves to an excellent interlinking hub in the LOD-Cloud [3].

This fast growing is due to the close bonds between DBpedia and Wikipedia. While Wikipedia grows fast through the effort of its large community, DBpedia also benefits from this new data. Every new or updated article in Wikipedia can be used to extract structured data for the semantic database with help of the DBpedia knowledge extraction framework [4]. In January 2011 the DBpedia datasets consist of over 3.5 million entities which are connected by over 672 million RDF-Triples [5].

While the search-capability of Wikipedia is limited to full-text search, the semantic database DBpedia offers new ways to find useful information. The Faceted Wikipedia Search[5] Application, which builds upon the DBpedia database, can handle semantic search queries like: ‘Actor of the movie The Lord of the Rings born in Berlin’, by using the semantic structure of the data. This enables the user to specify his search

[1] http://esw.w3.org/SweoIG/TaskForces/CommunityProjects/LinkingOpenData
[2] http://dbpedia.org
[3] http://www.wikipedia.org
[4] http://www.freebase.com
[5] http://dbpedia.neofonie.de/browse

query to find exactly the information he is looking for without browsing through hundreds of search results.

Freebase is a collaboratively created open database, to structure information from general human knowledge collected by a large community. Unlike DBpedia, users can directly edit and change the data and the structuring schema [6]. This means users by themselves can develop new types, categories and overall domains to structure the data they supply. This freedom led to a large community and accordingly to an enormous grows of thousands of new facts a week. Freebase was founded by Metaweb Technologies, Inc. in 2007 and consists already of over 20 million entities[6]. 2010 Google acquired Metaweb and the Freebase-Project to improve Google-Search with better support of semantic user queries by taking advantage of the freebase semantic dataset [7].

Overview of LOD-Datasets. In the following table we show an overview of well-known semantic datasets in the LOD-Cloud. The data shows a comparison in amount of data as RDF-Triples, the interlinking (both in and out coming links) and the domain of the datasets. Even though DBpedia does not include the most data, you can clearly see that it takes the major role in the interlinking process. This table also shows how important LOD became and how much data has been collected already.

Table 1. Overview of LOD-Databases (2009-2010) [8, 9]. Interlinking: low >10.000, medium >1 million, central >1 billion extern RDF-links.

Databases	Amount of Data	Interlinking	Domain
LinkedGeoData	> 3000 million	low	geographic
DBpedia	> 400 million	central	multi-domain
GeneID	> 170 million	medium	bioinformatics
Freebase	> 100 million	medium	multi-domain
RDF Book Mashup	> 100 million	low	literature
Geonames	> 90 million	medium	geographic
Musicbrainz	> 60 million	medium	music

2.2 Semantics Visualization

Semantic visualization is not a new field, but dealing with the massive amount of semantic structure on the Web of Data raises new demands for visualizing the structure of the data, while facing the problems of information overload. In this section we introduce two semantic visualization systems, which both use the approach of combining different visualization techniques to improve overview of structure, relations and detail of the data and give a short introduction in our previous work of the combination of visualizations in a knowledge Cockpit [10, 11].

Knoocks: (Knowledge Blocks) [12, 13, 14] is a desktop-application for the visualization of ontologies, which combines several visualization techniques for semantic data. To display concepts of a specific ontology, Knoocks uses the so called 'Knowledge Block' approach. A block is used to visualize a concept and includes all of its

[6] According to the Freebase Explore-Page: http://www.freebase.com accessed: 01/2011.

subconcepts within the block-boundaries. Besides the block-visualization of the structure, you can also use the outlook-window and a node-linked visualization component to view the semantic relation of the data.

Thinkbase: [15, 16] is a visualization and exploration tool for the Freebase database, which uses the Freebase UI as part of its interface. It consists of two combined views of the same data, a graph-visualization of the relational structure and the text-based Freebase interface, which shows detailed properties of the topic of interest. As can be seen in figure 1, the graph-visualization window shows all instances, which are directly related to the topic of interest and combines all relations with the same role into so called ‘aggregation nodes’ [15] to provide a better overview. The second part of the interface is the Freebase-UI, which contains textual information, images and other properties and relations of the selected topic.

Also this combined visualization has some advantages in giving the user the possibility to understand the neighboring structure of the instance in focus and simultaneously provide information about its properties, the systems deals with problems in adaptability and lucidity. The user cannot choose to see more than the direct neighbors of the instance in focus and every time he navigates through the graph, it changes completely, so he has to adapt to the new structure in every step. Also the Freebase-UI is not designed to be a part of a combined visualization, so it shows too much data, needs a lot of scrolling and is not interacting well with the graph visualization.

Visualization Cockpit. As described in our previous work, the visualization Cockpit [10, 11] is a user-adaptable approach of visualization combination. It lets the user combine different visualization techniques and representation details (like color, icons, depth of relational structure) into a personal ‘knowledge Cockpit’. Thus he can display the same data in different visualization techniques, which focus on different aspects of the data to improve a deeper understanding of the whole semantics structure. The user may for example combine the SeMap-Visualization [17] with a graph-like visualization to combine the advantages of the quick and easy overview capability of SeMap with the more detailed, but complex visualization of the graph.

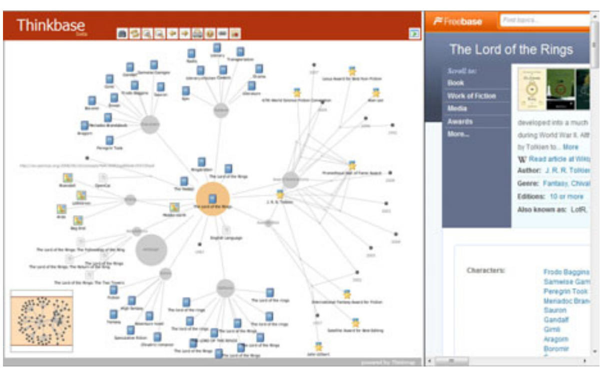


Fig. 1. Thinkbase visualization of the topic ‘The Lords of the Rings’

3 Conceptual Model

After introducing related work on visualization of LOD, we will describe our approach to overcome some of the disadvantages of these technologies. First we will describe our approach to systematic classify search results in LOD-Databases. Supplementary we will point out the weak points of common visualization techniques and finally introduce the idea of combining visualizations according to the underlying data, which varies with the results of user queries.

3.1 Extraction and Organization of Linked Open Data

The semantic structure of linked open data is that of an ontology, which consist of concepts, sub-concepts and instances, commonly available through RDF-Triples. This structure provides useful information in grouping the data into categories and sub-categories. Our approach includes the improvement of user search by making use of this preexisting categorization, to help users define their search query more precisely.

When a user defines a search query, most of the time he uses only one or a few words. The result commonly consists of a large set of instances, which correspond to the users search query with different relevance. The common view of the user is a list of results sorted by relevance, visualized by their names, maybe a short description or picture. With help of semantics data, we can improve the visualization of search results to help the user to further specialize his query. The existing categorization of the data can be used to structure the search results within these categories, so that the user can choose high order categories to decrease the number of relevant search results.

To make use of the semantic categorization, the system has to perform a number of queries against the database, which are sketched out in Figure 2. First the user query is passed to the LOD-Database to get a number of relevant instances (likely this number will be very large due to the massive amount of data). If the number of instances is too big to show in a list, the system performs a query to get ancestor-concepts of these instances. This procedure can be repeated until a reasonable number of categories are found, a breakpoint is reached or the hierarchical structure ends. We will end up having a set of categories which cluster the search results in reasonable sized groups and support the user in finding the information he is looking for.

3.2 Visualization of Semantics Data

Semantics data consist of a set of different and varying components. We have a hierarchical structure, which we used in the previous section to generate a categorization of search results. The instances have relations between each other and contain additional information in form of properties. These properties consist of varying data types, which includes number, date, time, geological features, text and so on. Knowing this, it becomes clear that a visualization of all these different features will be a different task to perform.

Most visualization techniques specialize upon one feature of the above. This is because the visualization methods they use have advantages for a special feature, but disadvantages for others. We can easily show the relations between instances in

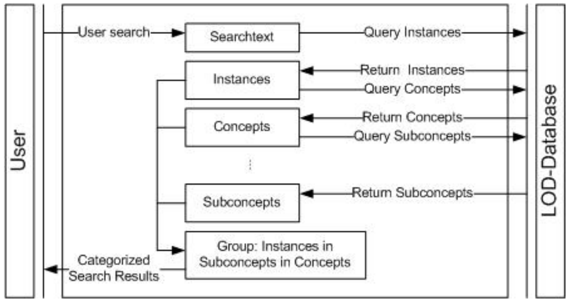


Fig. 2. Extraction of semantic data classification for search results

a graph-visualization, which provides interaction methods for expanding or collapsing a node to gain a better overview, but we can hardly display an article text, a picture or properties like geological or time related data. On the other hand can a map-based visualization be very comfortable to compare geological features of semantic data, but you cannot show the hierarchical structure of the data.

This raises the idea of combining different visualization techniques, as we pointed out in our previous work [10, 11]. To show all components of semantics data, a combination of visualizations is used to overcome the disadvantages of the individual systems by letting the user choose to look at the most suitable visualization for his means. Thinkbase also uses this promising approach by combining a graph-like view to show the structure of the data side by side with the Freebase-UI, to bring up the textual components. But there is still room for improvement. The graph-based view shows only the relations between instances, but cannot display the concept-hierarchy of the data, which can provide useful information in clustering and categorizing data. Also the animation of the graph is not suitable to let the user follow his navigation. As we mentioned before, the graph shows only the direct neighbors of an object and cannot vary the depth as the user likes. It also does not have additional visualization methods the user could choose from and the Freebase-UI is much too crowded to be one part of a combined visualization tool.

3.3 User Centered Adaption Based on Semantics Data

Our main approach is to combine the ideas of categorizing search results and combined visualizations for a more suitable representation of semantics data. As we pointed out in the previous sections, Thinkbase and Knoocks both follow the approach of having different visualizations combined. But both only provide static, predefined visualizations, which should cover the most common datasets. In our previous work we described the idea of letting the user select his combination of visualizations directly by himself [10, 11]. This approach requires a certain kind of understanding in the underlying data and may not be suitable for inexperienced users. The main improvement of our approach is the automatic, dynamic selection and combination of visualizations, which are most suitable for the dataset currently in focus. This means the system has to analyze the data, structure it and choose a promising set of visualization techniques, which present the elements of the data they are adapted to.

The first step to do this is structure the data which returns as result from a user query. In the previous section we showed a method of categorizing this data into their concepts and sub-concepts to provide a better overview of the returned results. But semantics data contain more than only hierarchical order. Semantic relations and properties of the instance in focus can easily be queried on demand.

After the data is structured, the system has to decide on the visualization composition. Following the processing of the user query, we give an overview of the returned instances. The visualization of the search results is dependent on the number of instances, the number of categories and the number of instances grouped in these categories. Commonly there will be a lot of instances which respond to the search query. In this case, an unstructured list of all results would be a poor choice. The system will instead choose to categorize the instances as we described above and add a simple hierarchical visualization like SeMap [17], which will provide easy to use and understand access to the categorized search results. In contrary if there are only a few instances, there is no need to further categorize the data and it can be shown as an ordered list.

The next step is to visualize the instance in focus. Before any user interaction, this will be the most relevant instance from the search results, which in many cases is the one the user is looking for. The system first has to analyze the information contained in the instance of focus, which could be number of semantic relations or the type of properties it contains. As an example the instance could be a famous person, so it may contain properties like age or weight, which will be displayed using a tabular structure. It may also contain textual descriptions, pictures or links to the Wikipedia entry, visualized in a textual content window. Time dependent data like birthday or death will be displayed in a timeline-view and any geographical data in a map-based visualization. The system will analyze all this information and combines the different visualization techniques into one (further adaptable) knowledge cockpit. According to the number of different visualizations the system may choose to display only the most relevant views and hide the others, giving the user the possibility to expand the visualization he needs.

4 Case Study

In this section, we will present a comparative view of an example search query and user interaction in Thinkbase and our data-adaptive system. We will show the advantages of our approach analyze the differences between the two systems and discuss the choices our system made to display data returned from the search query.

We start with the search query: ‘Moon’ with the intention to find out something about the novel ‘New Moon’ (assuming the user doesn’t know the correct title). Thinkbase first shows a list of the first 10 instances which correspond to the search query, ordered by relevance and displayed with their names and domain. The user has to scan through the list and expand it two times to find ‘New Moon (Book)’. In our system, the search results are categorized into not more than 10 domains, from which one is Literature and displayed using SeMap [17]. If the user expands Literature, he finds the instance ‘New Moon’ within a short list of books with moon in title (sorted by relevance).

The next step is the selection of the instance ‘New Moon’. Thinkbase builds up a graph-view of directly related instances and shows the ‘aggregation nodes’: Characters, Genres, Editions, Subjects and Webpage. You can also extract the

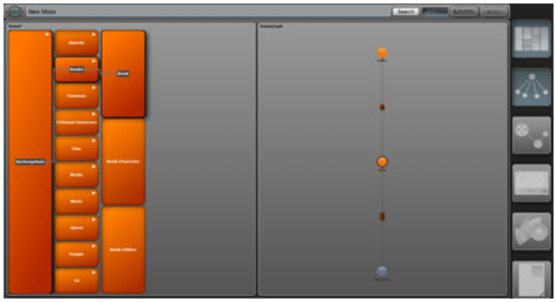


Fig. 3. Adaptive Visualization of Search Result

author, publication date and series the book belongs to. The Freebase-UI shows a short text and picture and contains as well all information, which is also part of the graph visualization (e.g. author, characters, editions …). Our data-adaptive system displays a similar graph-visualization to show the related objects, but offers the user the possibility to adjust the depth of the graph. It also adds a timeline-visualization, to show the publication date of the book alongside with the publication dates of the related books and release dates of the related films. Additionally, a text-based visualization shows an article text, a picture and links to Wikipedia.

In this step, you can clearly see the main difference between the two systems. While Thinkbase displays the same information in two different views, duplicating the majority of the data, our system clearly separates the information, displaying only the data the specific visualization is adapted to and combines them to an overall view of the semantic information.

5 Conclusion

Considering the fast development of the Web of Data it is clear that semantic annotated data is already an important part of the internet. While the amount of open data is steadily growing, the user interfaces still cannot reflect the whole power of semantics data. New visualization techniques are needed, to take advantage of the information, contained in semantic relations and structure.

In this paper we improved our previous work on user-centered combination of different visualization techniques to a new visualization system, which analyses, structures and visualizes semantic data automatically, according to the structure of the underlying data. We pointed out how search results can be categorized according to the preexisting hierarchical structure in Linked Open Data and how this could improve the information acquiring process of common users. To reduce the amount of information in a single visualization, the data is clearly separated into pieces and visualized by a combination of data-adaptive selected visualization methods.

To improve the user-centered design, further work should include user models, to adapt the combination of visualizations to users liking.

Acknowledgments. The here described work was developed as a part of the Core-Technology-Cluster for Innovative User Interfaces and Visualizations of the THESEUS Program, a 60-month program partially funded by the German Federal Ministry of Economics and Technology.

References

1. Bizer, C., Heath, T., Berners-Lee, T.: Linked Data - The Story so Far. International Journal on Semantic Web and Information Systems 5(3), 1–22 (2009)
2. Berners-Lee, T.: Linked Data. International Journal on Semantic Web and Information Systems 4(2), W3C (2006)
3. Auer, S., Bizer, C., Kobilarov, G., Lehmann, J., Cyganiak, R., Ives, Z.: DBpedia: A Nucleus for a Web of Open Data. In: The Semantic Web, pp. 722–735. Springer, Heidelberg (2007)
4. Bizer, C., Lehmann, J., Kobilarov, G., Auer, S., Becker, C., Cyganiak, R., Hellmann, S.: DBpedia - A crystallization point for the Web of Data. In: Web Semantics: Science, Services and Agents on the World Wide Web. Elsevier, Amsterdam (2009)
5. Bizer, C.: DBpedia 3.6 released (2011), `http://blog.dbpedia.org/2011/01/17/dbpedia-36-released` (online accessed: January 20, 2011)
6. Bollacker, K., Tufts, P., Pierce, T., Cook, R.: A Platform for Scalable, Collaborative, Structured Information Integration. In: Sixth International Workshop on Information Integration on the Web, AAAI (2007)
7. Menzel, J.: Deeper understanding with Metaweb (2010) `http://googleblog.blogspot.com/2010/07/deeper-understanding-with-metaweb.html` (online accessed: January 24, 2011)
8. Statistics on links between Data sets (2009), `http://esw.w3.org/TaskForces/CommunityProjects/LinkingOpenData/DataSets/LinkStatistics` (online accessed January 26, 2011)
9. Statistics on Data sets, `http://esw.w3.org/TaskForces/CommunityProjects/LinkingOpenData/DataSets/Statistics` (online accessed January 26, 2011)
10. Nazemi, K., Breyer, M., Burkhardt, D., Fellner, D.W.: Visualization Cockpit: Orchestration of Multiple Visualizations for Knowledge-Exploration. International Journal of Advanced Corporate Learning (IJAC) 3(4) (2010)
11. Nazemi, K., Breyer, M., Burkhardt, D., Fellner, D.W.: Semantic Visualization Cockpit: Adaptable Composition of Semanitcs-Visualization Techniques for Knowledge-Exploration. In: Conference ICL 2010 (2010)
12. Kriglstein, S., Motschnig-Pitrik, R.: Knoocks: New Visualization Approach for Ontologies. In: 12th International Conference on Information Visualisation, IV 2008, pp. 163–168. IEEE, Los Alamitos (2008)
13. Kriglstein, S., Wallner, G.: Knoocks - A Visualization Approach for OWL Lite Ontologies. In: 2010 International Conference on Complex, Intelligent and Software Intensive Systems (CISIS), pp. 950–955. IEEE, Los Alamitos (2010)
14. Kriglstein, S.: OWL Ontology Visualization: Graphical Representations of Properties on the Instance Level. In: 2010 14th International Conference on Information Visualisation (IV), pp. 92–97 (2010)
15. Hirsch, C., Grundy, J., Hosking, J.: Thinkbase: A Visual Semantic Wiki. In: Demo Session of the 7th International Semantic Web Conference (2008)
16. Hirsch, C., Hosking, J., Grundy, J.: Thinkbase: Interactive Visualization Tools for exploring the Semantic Graph of large Knowledge Spaces. In: Workshop on Visual Interfaces to the Social and the Semantic Web, VISSW 2009 (2009)
17. Nazemi, K., Breyer, M., Hornung, C.: SeMap: A concept for the visualization of semantics as maps. In: Stephanidis, C. (ed.) UAHCI 2009. LNCS, vol. 5616, pp. 83–91. Springer, Heidelberg (2009)

Riding the Technology Wave: Effective Dashboard Data Visualization

Lisa Pappas and Lisa Whitman

SAS Institute, Inc., 701 SAS Campus Drive, T31, Cary, NC 27513, USA
{Lisa.Pappas,Lisa.Whitman}@sas.com

Abstract. Riding the technology wave, we are awash in data. Attempts to stem the tide, or at least to manage its flow, have led to a proliferation of dashboards. With data dashboards, organizations consolidate important data in a single place, typically accessed via web browser. Dashboard contents may be tables, graphics, or visual key performance indicators (KPIs). While dashboards proliferate, displaying actionable data to support decisions, they are often developed by technical professionals inexperienced in human-computer interaction design. Research abounds on visual perception, but typically this is in the context of individual entities rather than composites. This research will survey corporate and institutional dashboards and their use of particular methods of data visualization. Building on existing research into effectiveness of interactivity in diagrams, the authors seek to identify which types of data are best represented in what types of visualizations and best practices for displaying multiple visualizations in a single view.

1 Introduction to Data Dashboards for Business Guidance

With ever-increasing data volumes and technology accelerating the speed of commerce, business leaders struggle to distill information into actionable intelligence. To make fact-based decisions, they need the right data, delivered reliably, in an easily accessed and perceivable form. The most typical form for this is the dashboard, a metaphor adopted from automobile or aircraft designs. But this metaphor – itself based on current situational data, such as speed and miles traveled – is limited to a snapshot of separate measures.

Although current status is important, particularly for operational data, decision makers require data in context to truly manage performance over time,. They require comparisons of current values to past performance and to future objectives. Time horizon and scope of data needed differ significantly based on roles in the organization. An executive, focused on achieving enterprise-wide strategic goals, requires a high-level view across lines of business and covering months or years. Line-of-business managers, tasked with meeting tangible daily or weekly performance goals, require not only a narrower timeframe and band of data, but also, if current rates are off-target, the ability to quickly dig beneath the surface to investigate the amount and cause of variance. Business analysts have a much broader set of needs. Rather than knowing what they are looking for, they often approach performance data with as yet unformed questions. With an understanding of corporate strategy, business analysts may combine the firm's

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 249–258, 2011.

operational and transactional data with external data sources, such as market demographics.

In addition to time and purpose limitations, traditional data dashboards fail to leverage human perception capability to facilitate swift understanding of status and intuitive guided analysis. Factors such as placement, attention cues, cognitive load, and interactivity contribute greatly to the effectiveness of a dashboard and the ultimate value it provides to its consumers.

In this paper, we identify three categories of business data dashboards and describe their dimensions and user base. Then, we discuss elements of dashboard design, including types of visuals and interactions available, and provide guidance on their situational utility. We conclude with a discussion of further research needed, particularly driven by the rapid spread of mobile devices as a physical platform for decision support.

2 Data Dashboard Categories

To be a successful business tool, a dashboard must "have a direct relevance to critical business activities" [1]. It must present data efficiently and provide a means for the user to take action based on the data presented. However, critical business activities and possible actions to take vary by user role in an organization. To determine what makes dashboard visuals effective, we must understand who is using them and for what purpose. Determining the intended recipient, we can surmise the level of currency required and the types of measures or categories relevant for the role. From the purpose, we can derive the data type, quantitative or qualitative, and scope, enterprise or departmental, most likely to be presented. Among other sources, Stephen Few's *Information Dashboard Design* [2] provides guidance for a useful taxonomy, proposing three high level categories: strategic, operational, and analytical.

2.1 Strategic

The most widely recognized use of digital dashboards is that of the executive dashboard. Its purpose is to communicate to management the organization's performance relative to corporate objectives. Its nature invites comparative data, contrasting current with past performance or current to target levels.

> *The strategic dashboard allows for a quick overview of an organization's 'health,' so to speak; assisting with executive decisions such as the formation of long-term goals. The strategic dashboard, therefore, doesn't require real-time data: what is going on right now is not important, what is pressing is what has been going on.* [3]

Strategic dashboard data may be quantitative or qualitative. Consider the concerns of an airline executive, with the dashboard shown in Figure 1. Quantitatively, the executive wants trend data on revenues, expenses, and profits. Qualitatively, s/he wants top 10 best and worst routes (based on cancellations and delays) and information that highlights successes and problem areas.

Because of their broad time horizon yet specific targets, strategic dashboards should have an uncluttered interface, to quickly guide consumers to the answers they seek. Simply put, "Are we on track?" While such dashboards are primarily provided for senior management, they are increasingly shared to a wider audience. Providing insights on performance across the enterprise can promote alignment through lines of business toward the corporate goals.

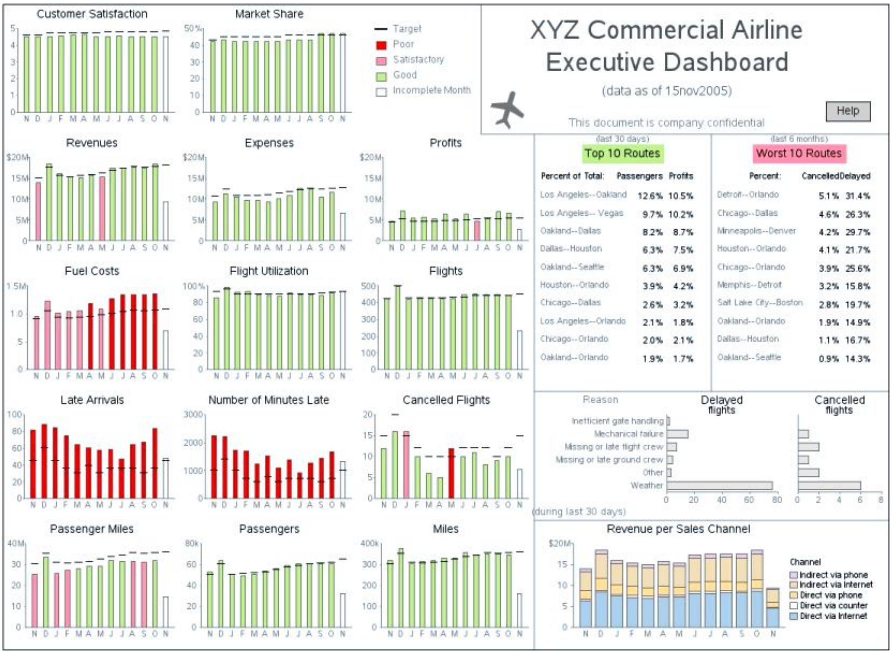


Fig. 1. Strategic Dashboard Example – Commercial Airline[1]

2.2 Operational

Monitoring operations requires more timely data, tracking constantly changing activities that could require immediate attention. Like strategic dashboards, effective operational dashboards require an uncomplicated view to enable rapid visual parsing to identify measures that are off-target and require intervention.

As with strategic dashboards, the display media on operational dashboards must be very simple. In the stressful event of an emergency that requires an immediate response, the meaning of the situation and the appropriate responses must be extremely clear and simple, or mistakes will be made [2].

Demand for currency of operational data can vary. As long as things are on track, periodic snap-shots may be sufficient. However, if a measure goes out-of-band, operational managers may want real-time data to see if the variance is an anomaly or a trend. In addition to a high level view, the operational dashboard must enable a

[1] Source: Used with permission of Robert Allison.

deep-dive, so that the business user can examine why a measure is off-target and take appropriate action. For example, the facilities department for a university might monitor volume of water usage, as a sudden spike could suggest a leak. By drilling down on a metering graph, facilities could isolate the usage anomaly geographically.

2.3 Analytical

Analytical dashboards share attributes of both strategic and operational dashboards. Like the strategic dashboards, the timeframes may be wider. Like the operational dashboards, drill-down and visual exploration are essential for discovering patterns and trends in the data.

> *More sophisticated display media are often useful for the analyst who must examine complex data and relationships….Analytical dashboards should support interactions with the data, such as drilling down into the underlying details, to enable the exploration needed to make sense of it – that is, not just to see what is going on but to examine the causes* [2].

In this regard, an analytical dashboard may exist at the intersection of strategic and operational data. For example, consider an organization conducting clinical trials for a new drug to treat diabetes. Healthcare providers at a number of sites around a metropolitan area record health statistics of trial participants at regular intervals, the operational data. But suppose that the trial sponsor, from the over-arching strategic dashboard, notices less favorable results in one particular region. The sponsor might then request a business analyst to examine the data to uncover the source of the variance. The analyst starts with regional data, but then slices the data by age, reported diet, or other variables to uncover relations between the data. For example, are participants in one region older than another, or are they more likely to smoke?

Besides looking back to examine root cause, analytical dashboards can be forward looking as well, helping to forecast outcomes. By examining current trends, business analysts can model outcomes by adjusting variables to recommend actions to optimize results.

Table 1. Summary of Dashboard Categories and Attributes

Category	Purpose	Timeframe for Insights	Data Scope	Update Frequency	Interactivity
Strategic	See and decide or question	Months or years	Enterprise-wide, cross-business unit	Moderate	Low
Operational	See and act	Minutes or days	Business-unit specific	High	Moderate
Analytical	See and question, explore what-if scenarios	Minutes to years	Enterprise-wide, cross-business unit, or isolated	Low	High

2.4 Summary

Table 1 summarizes the purpose and properties of the three different categories of dashboards discussed. With an understanding of the categories of dashboards, we now explore techniques for choosing appropriate dashboard contents for the different dashboard categories and best practices for designing dashboards that are usable and effective for various types of users based on their objectives.

3 Data Dashboard Contents and Design

A wide variety of visuals and interactions are available for use in computer-based dashboards. In order to design a dashboard that effectively meets the needs of its audience, one must take care to choose the right visuals and interactions based on their situational utility. A dashboard is meant to be viewed at-a-glance, so once the visuals have been selected, they must be arranged in a display that can be viewed all at once, such as a computer screen, without having to scroll or navigate to multiple pages. Information is effectively integrated, risk is quickly noticed, and decisions are most easily obtained when information is displayed using visuals that are arranged together so they can be seen simultaneously on one screen. This allows for processing the information with minimal effort [4]. In the summary table above, the first purpose for each of the three categories of dashboards is to "see"; therefore, it is crucial to design dashboards so that all important information is noticed quickly.

The audience and the data should guide your design. It is important to know who will be the users of the dashboard you are designing and what their goals are, so you understand which category of dashboard you will be designing. After obtaining this information through user interviews or requirements, you can determine what story the data need to tell on the dashboard, which messages are the highest priority (i.e., data that are monitored frequently or have the most critical consequences), how the data will be presented, and what actions the users should be able to take (i.e., interactivity with the dashboard, drill-down capabilities to obtain details, hyperlinks to additional or relevant information, and communication systems like commenting). To make these decisions you will need to communicate with the users of the dashboard. They need to know what you are capable of providing so that they can understand how much data are available to them, what different ways the data can be presented and technologies that can be used, and how data can help them make decisions and accomplish their goals. They need to communicate to you what information they need to know regularly and whether they monitor particular information frequently, where they go to view data (for example, a private computer, a shared computer monitor, or a mobile device), how they will want to view the data (for example, do they prefer quickly interpreted graphics or detailed tables), and what actions they will need to take (such as, making comparisons or drilling down for details). It is also useful to understand how your audience is accustomed to viewing data and with what visuals they are already familiar in order to choose visuals on your dashboard that will be easy for them to interpret.

3.1 Choosing the Right Visualizations for Your Dashboard

To design a dashboard that is effective and usable for your audience, you need to choose data visualizations that convey the message of the data clearly, are easy to interpret, avoid excessive use of space, are attractive and legible, guide the user to next steps (decide, question, or act), and allow the user to accomplish tasks to achieve their purpose (such as making comparisons or monitoring performance).

Choose visualizations that display data in a way that clearly and quickly enable users to achieve their purpose. For example, dashboards may provide the user with visualizations that compare data. Line graphs, bar charts, and bullet bars are effective visuals to use for quick comparisons. Strategic and analytical dashboards may provide interactivity, such as filtering or drill-down exploration, letting users further explore the causes of data variations noticed in the comparison. A scatter plot is another means to provide more detail behind comparisons by showing patterns created by individual data points, and bubble plots add an additional dimension in order to see what other variables might be causing variance. However, if the goal is to design an operational dashboard, then visuals showing comparisons should display any variations that would require action in a way that is quickly and easily noticeable. Making the visual a key performance indicator (KPI) is an effective way to show the comparison and also draw attention to data points that may require action. A KPI is set up to show where data falls within a specified range, so if a value falls below or above a threshold the visual utilizes color coding to draw attention to that value. Typically red is used to show when performance has fallen below a target, green indicates good performance, and yellow can be used to show that no action is required. If the audience of a dashboard is mixed or unknown, then it can be optimal to provide visuals that allow quick comparison of values while also using KPI color coding, such as the examples in Figures 2 and 3. Figure 2 combines a bar chart (which allows a user to easily compare values) with performance thresholds that reveal which categories may need action or further exploration. Figure 3 shows bullet bars, which are KPIs that allow a user to quickly compare values, alongside text that provides additional detail. If multiple KPIs are used in a dashboard, make sure that color coding is used consistently for the different KPIs, so a user does not have to go through the extra work of deciphering different color codes for KPIs that have the same meaning. For example, use the same shade of red for all KPIs on your dashboard that show if a measure is performing below a threshold [5]. The bullet bars in this example are a form of sparkline. Sparklines are visuals that, while small enough to fit alongside text, have a resolution high enough and design simple enough to help convey a message quickly [4]. Sparklines use space efficiently and can be shown next to labels or in tables. Examples of effective sparklines include bullet bars (which are effective for making comparisons when aligned in the same orientation), line charts (which, even when small, can reveal trends), and icons (such as red down arrows and green up arrows). Keeping the graphics in close proximity to relevant text helps the user to see them as related [5].

Less efficient data visualizations include pie charts, speedometers, and dials. Because these visualizations are round, they take up a lot of space relative to the amount of information that they deliver. This is an inefficient use of space when all of the visuals on a dashboard need to be on one screen. However, use of these visualizations may be appropriate if they are meant to draw attention to an important measure that could require immediate action or have a significant consequence. Besides being large, pie charts are also inefficient because people have difficulty comparing angles. For example, a user can scan left to right along a bar chart and compare bar heights more easily than that person could notice angle variations in pie slices. For comparisons, it is most effective to use visuals that require the user to compare line lengths with a common baseline (such as a bar chart or bullet bar) and less effective to use visuals that show comparisons with angles, area, volume or color [5]. This should be taken into consideration when choosing KPIs (such as those shown in Figure 4) for your dashboard.

Fig. 2. A bar chart with reference lines and KPI color coding[2]

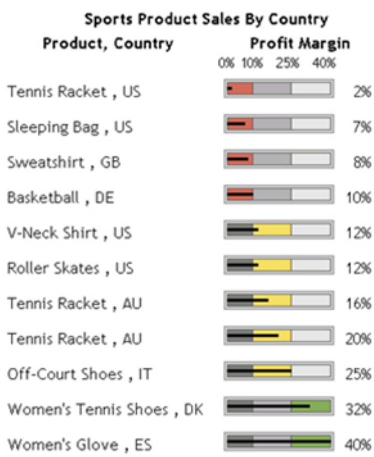


Fig. 3. Bullet bars[2]

You should use your dashboard real estate as efficiently as possible. If display space is limited, use efficient visuals like bar graphs or line graphs. If a lot of space is available, you can use less efficient visuals (like dials or speedometers). Pictures should be avoided, as they do not add as much value as a data visualization, and photographs of people are distracting since people tend to fixate on any faces present on a visual display. It is best to focus on the data and avoid any visual distractions. Background images should especially be avoided, since reading text laid over them is difficult. Motion attracts attention, so avoid using extraneous or looping animations, such as tickers. Using too many colors, or colors that are too bright, is also distracting [6]. Since some users may be color blind, color alone should not be used to convey meaning. Instead, combine color with intensity or border thickness. Use labels to show values. If colors are used to show comparisons in data, their shades should vary. Printing your dashboard with a black and white printer is a good way to test if the colors vary enough to be distinguishable even as different shades of grey [6]. When used appropriately, color can be effective for alerting the user of any "red flags." No other colors or graphics on the dashboard should distract from alerts or color coding. Extraneous text should also be avoided. A **visualization** should be easy to interpret without an explanation, so only important text (like graph titles, category labels, or data values) should be on the dashboard. While a dashboard may have a small area, text should not be made so small that it is difficult to read. Ask your user to view the dashboard where it will be displayed in order to test it. This will make you aware of any adjustments that need to be made to visualizations, terminology, fonts, or color. The user should see organized groups of relevant information. Visuals that are related to each other should be close to each other, with white space around the group. The users should be able to easily find important information. Since people read left to right, put the most important information in the top-left corner, or the information that a user will want to look at first. If an object is in the center of the display, it will be noticed first, so if anything is placed directly in the center of the display, it should be important [5].

Fig. 4. Various types of KPIs – a dial, a speedometer, a slider, and a bullet bar[2]

Interactions can help to show the user all the information they need, even in a limited amount of space. Tooltips that appear when a user hovers the mouse pointer over a graph can provide details and even links to more information. The ability to expand or collapse visuals can be useful when many visuals crowd a small display area, allowing users to focus on the information most relevant to them. Interactions like data brushing, or interactive highlighting, can also help a user to focus their attention on important data. Data brushing is a technique where, as the user changes data selection in one view, "corresponding linked data in one or more other views is highlighted" [7]. For example, they can click on a row in a table or move a slider along a line chart (as shown in Figure 5) to reveal a bar graph showing detailed information for the date or category selected. In this way, the user can choose what is displayed on their dashboard, so that space is utilized in a way that is most beneficial to them.

Fig. 5. An effective use of interaction for choosing information to be displayed[2]

4 Conclusions and Implications for Further Research

As described in previous sections, dashboard designers have many options when choosing methods to visualize data. But as with all effective design, they must let their design be guided by the users' goals, the context of use, and constraints – or lack thereof – of the underlying data. Availability of and reliance upon current, business critical data demands a broader skill set of designers. Rather than merely designing a visual display and specifying interactive behavior, the very task of designing is morphing and broadening. In *About Face 3*, Cooper, et al., propose that interface design is becoming the process of defining the product itself.

> *Properly deployed [design]…both identifies user requirements and defines a detailed plan for the behavior and appearance of products. In other words, design provides true product definition, based on goals of users, needs of business, and constraints of technology.*[8]

The authors' supposition is particularly relevant for dashboard design, especially as business users demand access to their data on a wider variety of devices with very different form factors. The dashboard landscape now encompasses everything from mouse-driven traditional desktop monitors to smaller-screened laptops with touch pad navigation. The real estate constraints of tablet computers and smartphones are offset with gestures for direct manipulation of data visuals. Returning to our dashboard categories, we suggest:

- For an executive management strategic dashboard, use efficient visuals that combine multiple indicators in one space and provide some interactivity, such as data brushing, for limited exploration of trends and data relationships. These may be effectively included on a range of devices.

---

[2] *Source of Figures 2-5:* SAS BI Dashboard, http://www.sas.com/technologies/bi/entbiserver

- For line-of-business managers who focus on a few truly key measures, a smartphone deployment may be feasible. The near-real-time currency of the data, persistent connection, and alert-and-notify capability can be achieved in the small mobile devices.
- The exploratory nature of analytical or forecasting dashboards requires greater real estate, although an effective dashboard could be delivered to a tablet, provided the data throughput and supporting infrastructure were sufficient. Consider a market analyst exploring stock price fluctuations timed the release of labor or pricing statistics; he is not certain where the data will take him, so providing optional views of data can be a helpful heuristic. For example, he could view a scatterplot, and noting a cluster of points, select and zoom to the cluster; then, he may want to view just the selected range in a table or bullet chart for alternative comparisons.

Designers of business interfaces face both new challenges and opportunities. Technical capability and choices are ever-expanding, as are expectations of business data consumers who want the information they need, when they need it, in an easy-to-perceive format, wherever they are. However, the opportunity thoroughly to research the need, use, and context to define a product makes the challenge worthwhile. Our understanding of human perception and learning can promote data dashboards whose elegant visuals are readily used and valued.

References

1. Fuchs, G.: Dashboard Best Practices, http://www.logixml.com/resources/whitepapers/04_Dashboards.pdf
2. Few, S.: Information Dashboard Design: the Effective Visual Communication of Data, pp. 40–46. O'Reilly Media, Inc, Sebastopol (2006)
3. Hetherington, V.: The Dashboard Demystified, http://www.dashboardinsight.com/articles/digital-dashboards/fundamentals/the-dashboard-demystified.aspx
4. Tufte, E.R.: Beautiful Evidence. Graphics Press, Cheshire (2006)
5. Wickens, C.D., Hollands, J.G.: Engineering Psychology and Human Performance. Prentice-Hall Inc., Upper Saddle River (2000)
6. Rasmussen, N., Chen, C.Y., Bansal, M.: Business Dashboards. John Wiley & Sons, Inc., Hoboken (2009)
7. Ward, M., Grinstein, G., Keim, D.: Interactive Data Visualization: Foundations, Techniques, and Applications, pp. 320–321. A. K. Peters, Ltd., Natick (2010)
8. Cooper, A., Reimann, R., Cronin, D.: About Face 3: The Essentials of Interaction Design, pp. 17–25. John Wiley & Sons, Indianapolis (2007)

Peer-to-Peer File Sharing Communication Detection Using Spherical SOM Visualization for Network Management

Satoshi Togawa[1], Kazuhide Kanenishi[2], and Yoneo Yano[3]

[1] Faculty of Management and Information Science, Shikoku University,
123-1 Furukawa Ojin-cho Tokushima, 771-1192 Japan
[2] Center for University Extention, The University of Tokushima,
1-1 Minami-Josanjima Tokushima, 770-8502, Japan
[3] Institute of Technology and Science, The University of Tokushima,
2-1 Minami-Josanjima Tokushima, 770-8506, Japan
doors@shikoku-u.ac.jp, marukin@cue.tokushima-u.ac.jp,
yano@is.tokushima-u.ac.jp

Abstract. In this research, we have built a system for network administrators that visualize the Peer-to-Peer (P2P) file sharing activities of network users. This system monitors network traffic and discerns traffic features using traffic mining. This system visualizes the P2P file sharing traffic activities of an organization by making the processing object not an individual user but a user group. The network administrator can comprehend the P2P sharing activities of the organization by referring to the map. This system extracts traffic features from captured IP packets that the users communicated. And this system extracts the appearance ratio of DNS host query. These extracts data are processed by MapReduce method for high-speed sorting and calculation. Afterwards this system creates traffic model. These features of the traffic model are emphasized by weighting. After that, the traffic model is visualized by High Speed Spherical Self-Organizing Map. This feature map shows network traffic behavior related with P2P file sharing communication like a birds-eye view. As a result, we think we can assist the monitoring operation and network administration.

Keywords: Traffic Mining, Traffic Visualization, Administrator Assistance, Peer-to-Peer communication Detection, Spherical SOM.

1 Introduction

Today, Peer-to-Peer (P2P) applications have become on the Internet. It is applied in the field of file sharing, voice over ip (VoIP) and some groupware application software. Especially, a lot of file sharing software has been designed on the P2P communication model. If Internet users want to get various kinds of data, the users can easily obtain various files and data using P2P file sharing application software. The file content often includes music CDs and video DVDs protected by copyright low, it is not appropriate to exchange these files. Moreover, popular P2P file sharing application software in Japan, such as Winny[1] and Share need a huge bandwidth

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 259–267, 2011.

because these applications send and receive large amounts of data. As a result, regular communication such as E-mail, Web browsing is obstructed by P2P applications traffic. Additionally, new P2P file sharing application software that is called Perfect-Dark is released for several years.

In Addition, a virus that causes data compromise has been generated for the P2P file sharing network. There are many cases of classified information being compromised because of these viruses. For example, it is called Antinny, these viruses and their varieties posed a lot of cases of data compromise problem. These viruses give out sensitive information stored in the user's computer using the publication function on the P2P file sharing application software. When an organization's member uses a P2P file sharing application, there is a risk that the organization's security will be compromised. If certain classified information leaks, the data will pass along from place to place in the P2P network.

For example. many data compromise accidents occurred in Japan between 2004 and now. In 2004, an investigation content of one prefectural police department was leaked by viruses. In 2005 and 2006, a large amount of military intelligence and investigate information were leaked from Japan Self-Defense Force's departments. In 2010, a large number of prescription and medication history with customer information were leaked from one pharmacy company. In addition, a lot of companies customers were leaked them private information, and these are continuing compromise incident today. These accidents are extremely serious situation. And, these company and organizations are easy to lose customer's trust.

There are limitation techniques which use packet filters to limit illegal traffic that deviates from the policy established by the company or university. Filter technology, which synchronizes with the packet filter definition, is installed on the boundary defense devices such as firewall. And, illegal traffic that does not conform to the site policy is not forwarded to users. However, it is difficult to keep the filter definition perfectly set because the default destination port of each P2P application is different. Moreover, P2P application such as Winny and Share select the destination port number dynamically. This P2P application's traffic limitation can be impossibly strict. Therefore, P2P traffic cannot be limited only packet filter technology based on packet filter definitions.

On the other hand, the network administrator can use a specialized firewall system to limit P2P traffic such as One Point Wall[3]. These P2P traffic limitation techniques are based on the signature information that is extracted from illegal traffic. It is completely blocked when a signature matches a traffic pattern. However, when a signature does not match the traffic, illegal traffic is not restricted. Accordingly, the network administrator has to understand the P2P application's behavior in the organization's network traffic. If the administrator can understand the P2P application's behavior, the administrator can usually ascertain problem at an early stage.

At the present time, if an administrator wants to understand P2P application activities, protocol analysis such as sniffer tools can be used. However, this method is very labor intensive, and these methods only provide basic information like as IP address and port number level classification summary. The network administrator really wants a result that shows where the P2P file sharing application is used?

For these reasons, we have developed a traffic visualization system for P2P communication detection and administrator assistance. This system provides a feature map of traffic behavior made up from results of network traffic mining. This system assists the monitoring operation of the administrator by showing the feature map that this system presents. As a result, we can assist the monitoring operation of the administrator.

We pay attention to the traffic that the organization users send and receive. These features are extracted from this traffic using traffic mining. The features are the source/destination IP addresses, and source/destination TCP (or UDP) port number and TCP flags condition. In addition, we pay attention to the results of DNS query from internal clients. We found that when the P2P nodes try to find other nodes, the DNS query amounts are less than normal DNS queries such as Web browsing. We use these DNS query features for discerning P2P application behavior. Moreover, this system acquires the packet occurrence frequency and yield. Consequently, a traffic model is generated from the feature and packet occurrence frequency and a result of the DNS query. And, these traffic mining processes are accelerated by MapReduce algorithm based on distributed computing infrastructure.

The method of generating the model is the Vector Space Model (VSM). The similarity problem between traffic features is replaced with a cosine measure between vectors. Weighting is added to the obtained traffic model to emphasize feature quantity. Afterwards, a feature map is generated by using High Speed Spherical Self-Organizing Map (HSS-SOM) from the traffic model. This method enables to make accurate clustering results, because HSS-SOM does not have the problem of 'Border Effect' compared to Flat SOM.

This map shows an administrator which computer communicated to other computers and the volume of the communication. It expresses not only the summarized traffic amount but also each traffic type and behavior. It can be said that the feature map is a result of traffic mining from the users' traffic, and the administrator is assisted in understanding organization's traffic behavior by this map.

In this paper, we proposed a system framework of traffic visualization for P2P communication detection, and we show a configuration of the prototype system. Next, we show the results of experimental use and examine these results. Finally, we describe future study, and we show conclusions.

2 Assisting the Detection of P2P File Sharing Traffic

In this section, we describe the framework of P2P file sharing detection, and we show the result of exploratory experimental and results for traffic feature extraction.

2.1 Framework of P2P File Sharing Detection

Fig.1 shows a framework of administrator assistance of P2P file sharing traffic detection. We assist the monitoring and detecting operation of the network administrator by providing the traffic behavior of the organization users.

We paid attention to traffic between internal side and the Internet. All users' traffic passes a gateway in the internal site. We collect all IP packets that pass the gateway. In addition, traffic features are extracted from collected IP packets. Moreover, we paid attention to the results of DNS host queries from internal DNS resolver to destination's DNS servers. In most cases, P2P nodes information is distributed without hostname

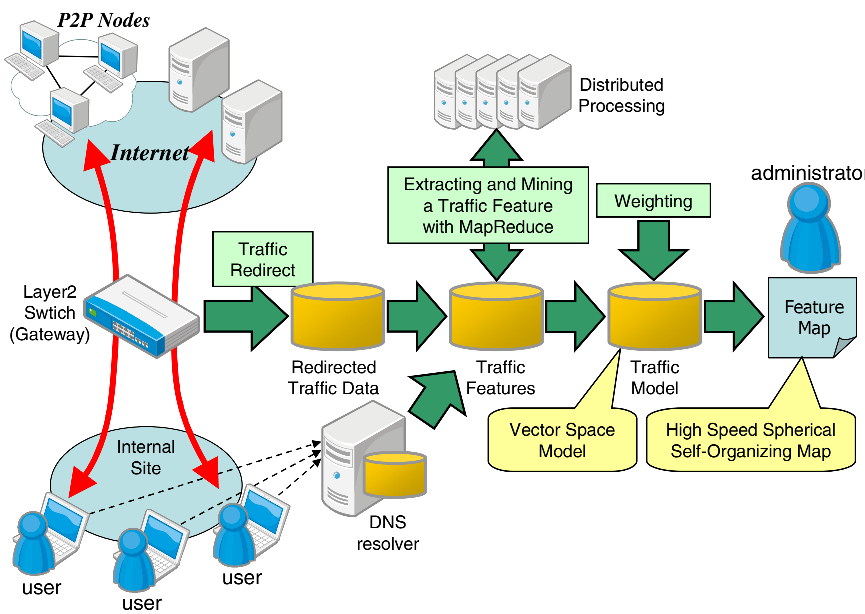


Fig. 1. Framework of administrator assistance

including FQDN (Fully Qualified Domain Name). In the result, the DNS host query amounts of the internal P2P nodes are less than normal applications. If the traffic feature of one host had different from other hosts, and that host's DNS host query amount was less than other hosts, that host has high probability of the P2P node.

Consequently, a traffic model is generated from extracted features that the users communicated and the results of DNS host queries. The method of generating the traffic model is Vector Space Model. As a result, the similarity problem between source IP addresses is replaced with a cosine measure between feature vectors. Moreover, weighting is added to the obtained traffic model to emphasize feature quantity.

A series of processing described here are traffic mining. Because the feature related to P2P file sharing communication is extracted from all captured traffic by the series of processing. Moreover, extracted and emphasized features are stored to the traffic model. This model adapt to traffic feature visualization.

Afterwards, a feature map is generated by High Speed Spherical Self-Organizing Map (HSS-SOM). HSS-SOM is an algorithm to map multi-dimensional vectors on a spherical space map. As a result, this map expresses the typical source IP addresses that the users communicated. The administrator gets birds-eye view of the organization communication activities by referring to the map. Therefore, the administrator is assisted in understanding P2P traffic behavior by this feature map.

2.2 Exploratory Experiment and Results for Traffic Feature Extraction

We made an exploratory experiment to clarify a feature of DNS host query by the P2P node. Especially, we want to clarify a DNS host query feature of the pure P2P nodes.

The P2P file sharing application was installed to experimental computer, and the experimental computer was used for 20 minutes with P2P application. After that, we generated general Web browsing traffic with other computers. Then both traffic was monitored and compared for this exploratory experiment.

Table 1 and Table 2 show the measuring results of the exploratory experiment. First of all, we can find a difference of the amount of sending IP packets. The case of pure P2P communication model much than Web browsing case per same measurement time. It is about 3.5 times larger than Web browsing case. That mean is P2P communication model makes a lot of a lot of connections between internal node and P2P nodes on the Internet.

Table 1. Measuring Results of pure P2P communication

Amount of sending IP packets	22,235
Amount of destination IP addresses	415
Appearance Ratio of TCP PUSH flag	35.2%
Appearance Ratio of DNS host query	Less than 0.1%

Table 2. Measuring Results of general Web browsing communication

Amount of sending IP packets	6,416
Amount of destination IP addresses	42
Appearance Ratio of TCP PUSH flag	9.2%
Appearance Ratio of DNS host query	91.8%

And, it is understood that appearance ratio of DNS host queries by P2P communication is remarkably low. In most situations, P2P nodes information is provided without hostname including FQDN (Fully Qualified Domain Name). As a result, it is provided only IP addresses. Therefore, DNS hostname resolution is not required to make connections between both P2P nodes. When the connection between P2P nodes is generated, the DNS host query is not appeared. Therefore, we can find striking difference of an appearance ratio of DNS host query between both communication models.

We can find small display of an appearance ratio of TCP PUSH flag between both communications. And this feature is variable in amount. When we use appearance ratio of TCP PUSH flag, we have fear of erroneous decision for detecting P2P communications. As a result, we think important features for detecting P2P communication are the appearance ratio of DNS host query and the amount of sending IP packets.

3 System Configuration

We show the configuration of proposed system in Fig.2. This system has 5 modules that include a "Traffic Collection Module", "Traffic Analysis Module", "DNS Query Analysis Module", "Modeling Module" and "Visualization Module". Especially, Modeling Module has a high-speed sorting and calculating function with MapReduce Algorithm, that works based on distributed computer system.

IP packets that users of organization sent and received are redirected via Layer2 switch with port mirroring function. Traffic Collection Module accepts the redirected IP packets from L2 switch. In addition, an Ethernet adapter configuration of this system is set to promiscuous mode. Because, this module have to accept all related IP packets. The accepted IP packets include normal traffic and illegal traffic, and all accepted IP packets are passed to the Traffic Analysis Module.

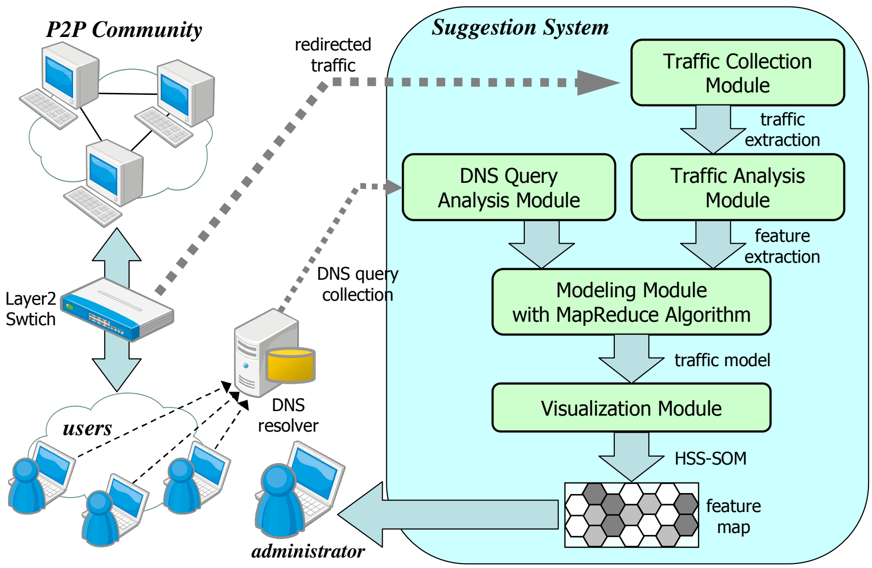


Fig. 2. System Configuration

Traffic Analysis Module attempts selection for all accepted IP packets. First of all, an administrator gives the IP address information of internal site servers to this module. And source traffic of the internal servers is dropped from all accepted traffic by using the given IP address information. Next, this module attempts to select traffic features from selected IP packets.

This module analyzes a packet field of selected IP packets, and some feature extracted from selected packets. The features are the source/destination IP address, and the source/destination TCP PORT number and TCP flags status. At the same time, each packets occurrence rate is calculated and stored. All extracted and calculated features are passed to the Modeling Module with other features generated from DNS Query Analysis Module.

The DNS resolver processes a DNS host resolution requests that was required from internal users. This module collects the results of DNS host resolution and requested client information from DNS resolver's log. It is selected excluding the incomplete results of DNS request. All extracted complete results of DNS hostname resolution are passed to the Modeling Module.

Modeling Module generates a traffic model which is defined by Vector Space Model. This module has MapReduce function that is pre-processing for generating a traffic model. An extracted feature data has a lot of same key-store value. This pre-process function makes a summarized data from a lot of same key-store value. After that, one source IP address corresponds to one multi-dimensionally composed vector, and each element of the multi-dimensional vector stores are number of destination IP address and the destination PORT number. We call this multi-dimensional vector a “feature vector”. The number of feature vectors is the same as the total amount of extracted source IP addresses. The set of these feature vectors becomes the traffic model.

The weighting process done to the feature vectors emphasizes the characteristics of the traffic model according to the occurrence rate with which the source IP address and DNS host query ratio. As a result, if the module discovers frequency appearing source IP address, it is possible to find the P2P packet spreader host.

Visualization Module visualizes and making the feature map from obtained traffic model. The High Speed Spherical Self-Organizing Map is used as a visualization method in this module. The source IP addresses of the processing object are self-organized by HSS-SOM algorithm. This results in a well-consolidated visual output that allows the administrator to get birds-eye view of internal users' P2P communication activities.

4 Experimental Use and Result

This system was tested to confirm its effectiveness. We collected traffic from users belonging to one organization at December 2010. The amount of observed data extracted from collected IP packets and generated feature vectors are presented in Table 3. And, Table 4 shows the computer's specification of experimental use.

Table 3. Amount of Observed Data

Data Type	Amounts
Observed Data	4,330,490
Generated Feature Vectors	39,632

Table 4. Specification of the Experimental Use

CPU	Intel Core 2 Quad Q9550
System Memory Capacity	8.0Gbytes
HDD Capacity	2.0Tbytes
Operating System	Microsoft Windows 7 64bit ed.

Feature maps were generated once a 25 minutes in the experimental period. The traffic data was put into the system. The input packet amount was about 9 million packets and the source IP address count that the system extracted was about 800. The feature map that the system generated had 80 elements, and each element corresponded to a summarized source IP addresses. Therefore, the source IP addresses appearing in the map where communicated a lot of times related with P2P communication.

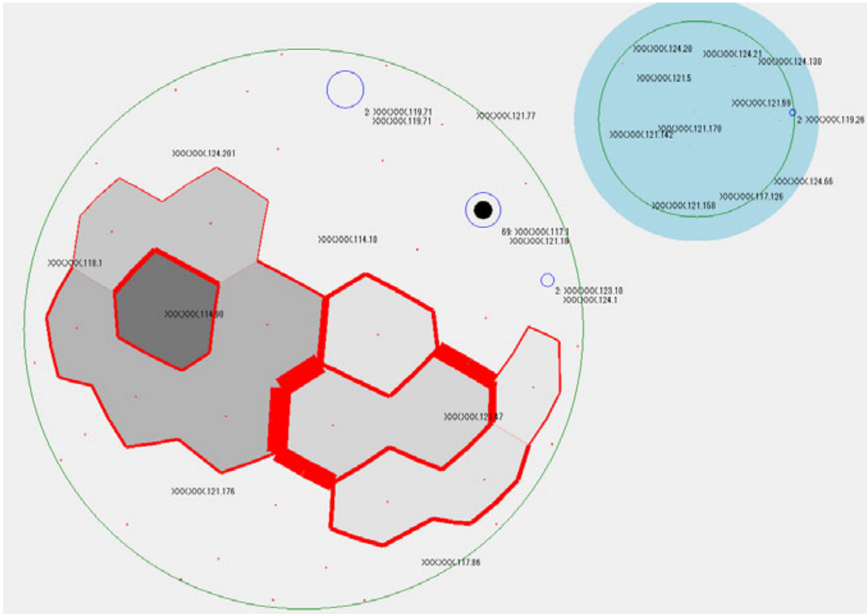


Fig. 3. Feature Map

We show generated feature map in Fig.3. We can find two large clusters on this map. One large and dark cluster declared the host of P2P file sharing communication.

At a later date, we checked traffic data deeply related to these cluster, we confirmed each host processed file sharing communication. As a result, we think that we can make complete to detect that file sharing communication. The network administrator assisted to detect the P2P file sharing traffic using this feature map and traffic mining method.

5 Conclusion

In this paper, we proposed a traffic mining system for P2P communication detection. And we explained a configuration of this prototype system. And, we shown the results of experimental use and examine.

This system extracts records of P2P communication activities from the collected IP packets and the collected DNS query results. In addition, this system provides a feature map for the administrator. We developed a prototype system and experimented to confirm its effectiveness. It was shown that an administrator could inspect the results of feature maps.

References

1. Kaneko, I.: The Technology of Winny. ASCII Books, Tokyo (2005)
2. BitTorrent Web Site, `http://www.bittorrent.com/`

3. One Point Wall, Net Agent Co, Ltd., http://www.onepointwall.jp/eng.html
4. Kohonen, T.: Self-Organizing Maps, 3rd edn. Springer, Heidelberg (2001)
5. Yoshida, K., Katsuno, S., Ano, S., Yamazaki, K., Tsuru, M.: Stream Mining for Network Management. IEICE Trans. Communication E89-B(6), 1774–1780 (2006)
6. Sugimoto, S., Tachibana, K.: A First Attempt to Apply High Speed Spherical Self-Organizing Map to Huge Climate Datasets. In: SOLA 2008, vol. 4, pp. 41–44 (2008)
7. Brenann, D., Hulle, M.: Comparison of FlatSOM with Spherical SOM: A Case Study, http://simone.neuro.kuleuven.be/marc/brennan.pdf

Visualizing Stakeholder Concerns with Anchored Map

Takanori Ugai

Fujitsu Laboratories Ltd.,
4-1-1 Kamikodanaka Nakaharaku Kawasaki 211-8588 Japan
ugai@jp.fujitsu.com

Abstract. Software development is a cooperative work by stakeholders. It is important for project managers and analysts to understand stakeholder concerns and to identify potential problems such as imbalance of stakeholders or lack of stakeholders. This paper presents a tool which visualizes the strength of stakeholders' interest of concern on two dimensional screens. The proposed tool generates an anchored map from an attributed goal graph by AGORA, which is an extended version of goal-oriented analysis methods. It has information on stakeholders' interest to concerns and its degree as the attributes of goals. Results from the case study are that (1) some concerns are not connected to any stakeholders and (2) a type of stakeholders is interested in different concerns each other. The results suggest that lack of stakeholders for the unconnected concerns and need that a type of stakeholders had better to unify their requirements. And a preliminary evaluation suggests that the tool enables users to identify imbalance of stakeholders or lack of stakeholders faster and more correctly than a matrix of stakeholders and concerns.

Keywords: Stakeholder management, Anchored map, visualization.

1 Introduction

Software development is a cooperative work by stakeholders. Because software is developed to be successful, it is extremely important to accurately understand user requirements. It is essential that the mutual agreement be handled efficiently and appropriately. However, if the stakeholders are missing, high-quality requirement specifications cannot be made, because knowledge is stakeholder-dependent. The case of the London ambulance service [7] is a typical case of missing stakeholders. It is important for project managers and analysts to understand stakeholder concerns and to identify potential problems such as imbalance of stakeholders or lack of stakeholders. Sharp at el [19] propose a technique to discovering all relevant stakeholders of a specific system.

This paper presents a tool which visualizes the relationships between stakeholders and their interest of concerns using an anchored map [14, 20] to detect the problems.

Anchored map is a technique to draw a bipartite graph whose nodes have two types of attributes. Anchored map fixes the nodes with one of the attributes on a concentric circle as anchor nodes and locates other nodes with a spring layout [6] as free nodes. The proposed tool generates an anchored map from an attributed goal graph by AGORA

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 268–277, 2011.

(Attributed Goal-Oriented Requirements Analysis) [10, 11]. AGORA is a family of goal-oriented requirements analysis [12] and Goal-oriented requirement analysis (GORA) is supposed to be one of the most promising methods for supporting the elicitation of requirements. GORA methods such as i* [3] and KAOS [5] are top-down methods for decomposing and refining customer needs into concrete goals to be achieved in order to satisfy customer needs. As a result, an AND/OR graph is generated whose nodes represent identified goals. There are a number of case studies and assessments of this method [2].

An AGORA goal graph is attributed AND-OR goal graph. A semantic tag is one of the attributes and abstracts of the meaning of the goal, such as a stereotype of UML. Stakeholder concerns are represented by semantic tags on the goals.

This tool will enable to support the Sharp's technique if the model of the technique is applied instead of the semantic tags of AGORA goals.

This paper is organized as follows. At first, AGORA and anchored map is described as basic technologies. And the way how to generate an anchored map from a goal graph is described. Then an integrated implementation of anchored map and AGORA editor is shown with a case study and a preliminary evaluation. At last, after discussion of related works, summery and future works are described.

2 Preliminaries

2.1 AGORA

AGORA [10, 11] is a part of the GORA family of analysis methods. To make up for the support functions that the existing GORA method does not have, AGORA uses an extended version of AND-OR goal graph called *attributed goal graph*. In AGORA, requirements analysts and stakeholders can attach *attributes* to nodes and edges of a graph. Fig.1 shows an excerpt example of the attributed goal graph. And Fig.2 shows

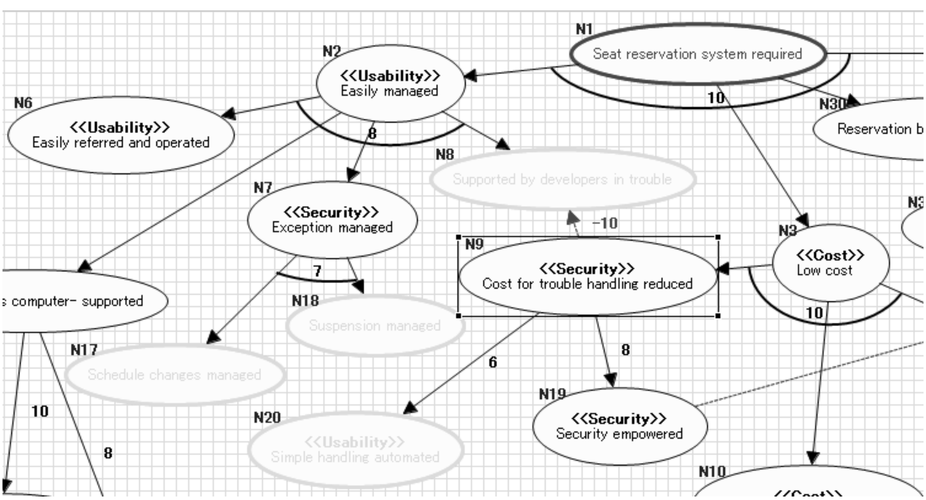


Fig. 1. An Attributed goal graph (excerpt)

snapshots of AGORA tool, including In these examples, the requirements of a seat reservation system for trains are analyzed. Although AGORA can deal with many types of attributes, we only use and explain the following two types of attributes in this paper:

- *Semantic tags*: Abstractions of the meaning of the goals, like stereotypes of UML. In the example above, the facilitator puts a semantic tag ``Security'' to a goal N_9 "Cost for trouble handling reduced" on a sub-window shown in Fig.2(a), representing that N_9 is related to security. We have extracted the possible candidates of semantic tags from, e.g., quality attributes enumerated in IEC/ISO 9126-1 [9] or NFR catalogs [4]. Requirements analysts can use additional semantic tags, e.g., domain-specific ones, if necessary.
- *Preference matrix*: Attached to a node, i.e. a goal, and represents the degree of preference or satisfaction of the goal for each stakeholder. Each preference is presented as a number from -10 to 10, and the higher value denotes that the goal is more preferable for a stakeholder. When the number is zero, he/she is not interested in the goal. Fig.2(b) is an example of the preference matrix, attached with the goal N_9: "Cost for trouble handling reduced". The columns, each of which is for the user (User), the administrator (Administrator), and the developer (Developer) are respectively to grade their own preference and preferences of other stakeholders. In this example, the values on the diagonal are -5, 9, and 0, and the analyst can understand that the user and the administrator have a conflict of interest at the goal N_{20}. The large variance of these preference values in a goal can suggest the conflicts of stakeholders' interests and the discordance of its understanding [11]. In this paper, we only use diagonal elements of the matrix, i.e., stakeholders' preferences by themselves.

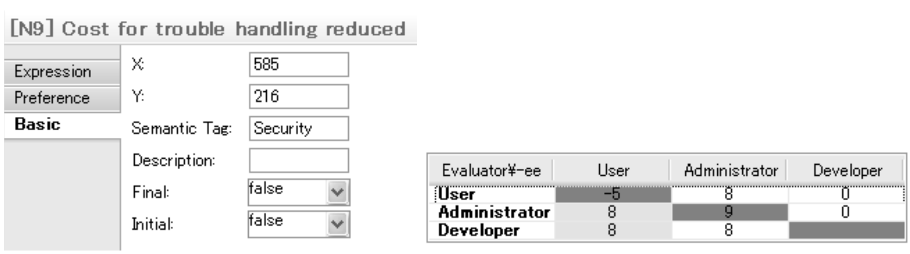


Fig. 2. Screenshots of AGORA tool: Left(a): Putting semantic tags for N_9, Right(b): Preference matrix for N_9

2.2 Anchored Map

This section presents the method to calculate the position of nodes in an anchored map for a bipartite graph. A bipartite graph consists of two sets of nodes and there are edges between nodes of different sets but there is no edge between nodes within a set. Anchored map is a layout technique for a bipartite graph.

Suppose set A consists of 9 nodes representing stakeholders which are User1, .., User3, Admin1, .., Admin3 and Dev1, .., Dev3 and set B consists of 8 nodes

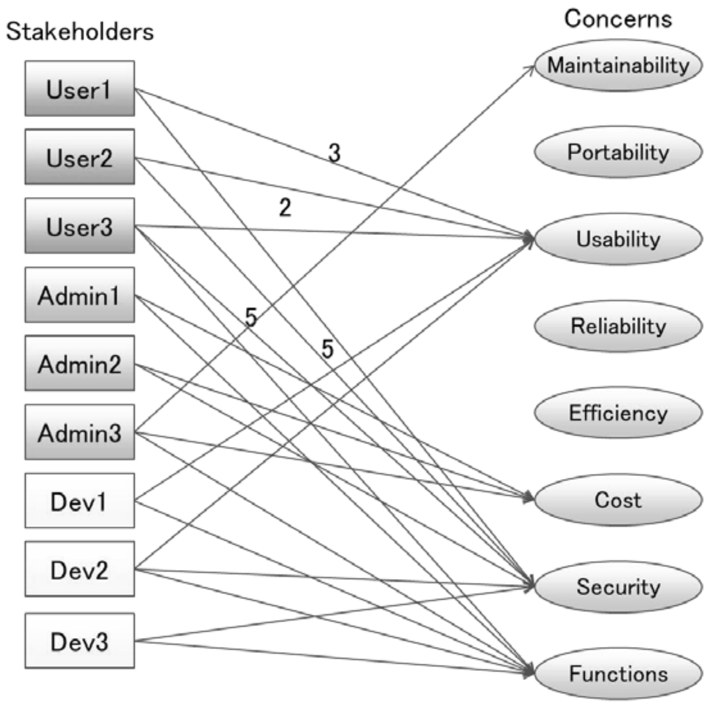


Fig. 3. Two-way layout for a biparate graph

representing concerns which are Maintainability, Portability, Usability, Reliability, Efficiency, Cost, Security and Functionality. Fig.3 shows a two-layer graphical representation of the same bipartite graph. Fig.4 shows the same bipartite graph shown in Fig.3 as an anchored map. On an anchored map of the bipartite graph, the nodes in A, called "anchor nodes", are arranged on the circumference and the nodes in B, called "free nodes", are arranged at suitable positions in relations to the adjacent anchor nodes. The strength of edge as a spring is in proportional to the value attached the edge. This means an anchor node and a free node are located closer to each other if the stakeholder is strongly interested in the concern. The distinction of the nodes is clearer than in Fig.3, and it is easy to see the relationships between free nodes.

In the case that supposing the elements in B as anchor nodes and the elements in A as free nodes, a anchored map is generated with the same way. This means two type of anchored maps are generated from a bipartite graph.

The printing area is 122 mm × 193 mm. The text should be justified to occupy the full line width, so that the right margin is not ragged, with words hyphenated as appropriate. Please fill pages so that the length of the text is no less than 180 mm, if possible.

3 Generating Anchored Map from a Goal Graph

Selected concerns are based on ISO/IEC9126-1 [9]. ISO/IEC9126-1 represents the latest research into characterizing software for the purposes of software quality control, software quality assurance and software process improvement. The ISO 9126-1 software quality model identifies 6 main quality characteristics, namely: Functionality, Reliability, Usability, Efficiency, Maintainability and Portability.

Cost is not one of the software quality factors in ISO/IEC9126-1 but is identified on many goals.

All goals are aggregated by the concerns represented by the semantics tags. The strength of the edge between a stakeholder and a concern is the average of the absolute value of the stakeholder's preferences of the aggregated goals by the concern. This means a stakeholder and a concern are located closer to each other if the stakeholder is strongly interested in the concern.

Because of this data representation, this tool enables the users to notice lacks of stakeholders to concerns. Anchor nodes are located on the circumference at equal intervals, and the positions of anchor nodes are decided as the number of crosses of edges is reduced as far as possible. Using this positioning algorithm, the free nodes that have a similar relationship with anchor nodes are located closer to each other. Because of this positioning algorithm, this tool enables the users to notice imbalance of stakeholders.

4 Implementation and Preliminary Experiments

4.1 Implementation

The tool is implemented as a plug-in of eclipse The input is produced through AGORA tool, and the tool extracts goals with attached preference matrix and semantic tags. The layout of an anchored map is implemented based on the algorithm shown in [14].

The tool is implemented as an integrated tool of anchored map and AGORA but the idea of visualizing relationships between stakeholders and their concerns does not depend on AGORA. User Requirements Notation (URN) [1] is a possible notation. Relationships between a URL model and actors can be an input of the tool.

4.2 Preliminary Experiments: A Case Study

Nine stakeholders put their preferences and the tags. These stakeholders consist of three end-users (User1, User2, User3), three developers (Dev1, Dev2, Dev3) and three system administrators (Admin1, Admin2, Admin3). In Fig.4, squares represent stakeholders and small circles represent concerns. Squares are color-coded according to stakeholder roles such as end-user, developer and administrator.

In Fig.5, stakeholders are located on a circle as anchors and concerns are arranged as free nodes. From the figure, we can see that,

- Most of stakeholders are strongly interested in security and functions because they are located in the center of the circle.
- Developers, Admin1 and Admin3 are interested in cost but users are not.
- Users and Admin2 are interested in usability but developers are not.
- Only Admin2 is interested in maintainability.

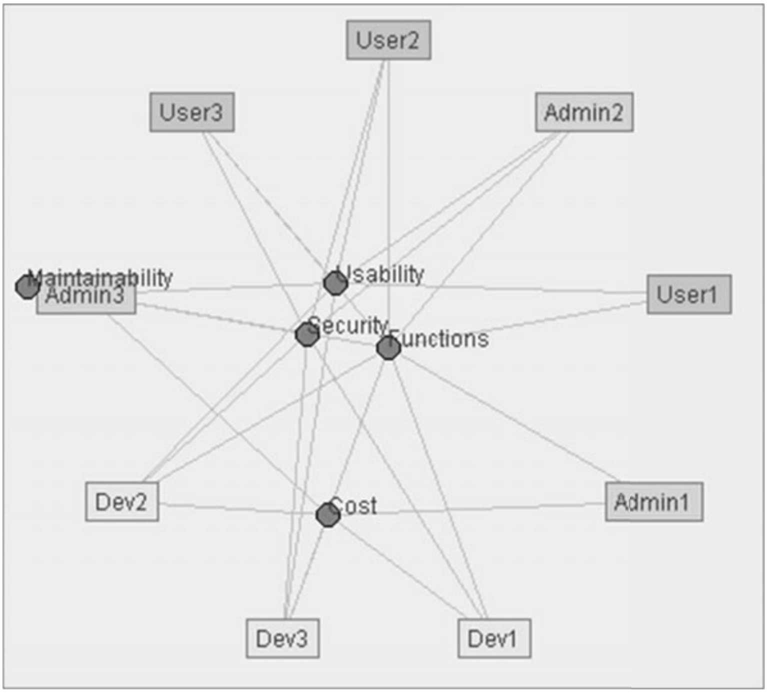


Fig. 4. Example of Anchored Map (Anchor nodes: stakeholders, Free nodes: concerns)

In Fig.5, concerns are located on a circle as anchors and stakeholders are arranged as free nodes. From the figure, we can see that

- All users are interested in similar concerns, and Admin2 is also the same as the users.
- Dev2 and Dev3 are interested in similar concerns but Dev1 is slightly different from Dev2 and Dev3.
- All administrators are interested in different concerns.
- Nobody is interested in portability, reliability and efficiency. This is because the goal graph describes business domains and does not describe the system infrastructure.

The findings suggest that (1) stakeholders who know much about the system infrastructure should be involved in the project and (2) differences of administrators should be consolidated.

The case study shows the system visualizes imbalance of stakeholders, lack of stakeholders, and some other potential useful information to manage stakeholders.

4.3 Preliminary Evaluation

We conducted experimentation whether the tool enables users to identify imbalance of stakeholders or lack of stakeholders faster and more correctly than an existing technique.

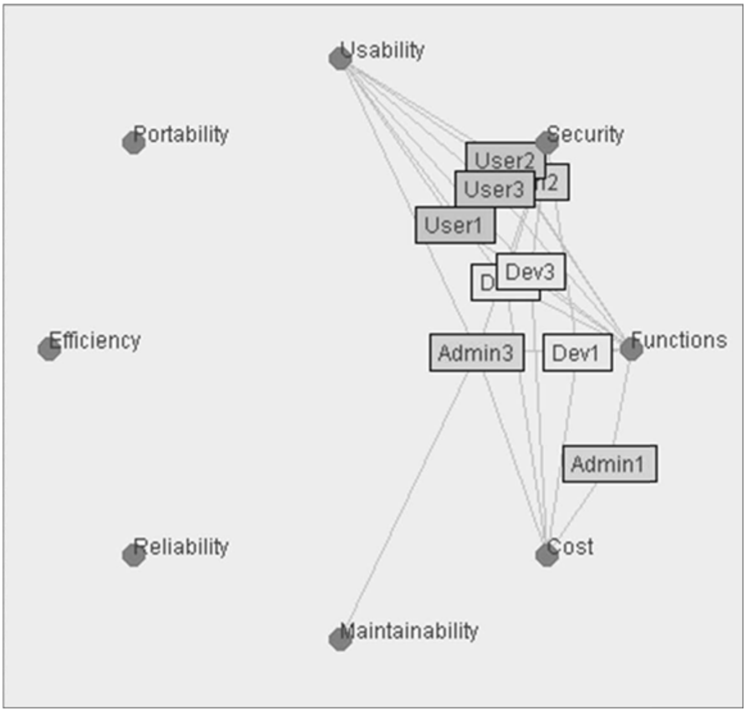


Fig. 5. Example of Anchored Map (Anchor nodes: concerns, Free nodes: stakeholders)

Table 1. Goal Graphs

System	RSS Reader	Reception
# of nodes	29	36
# of concerns	4	5
# of stakeholders	**9**	9

Preparation

- *Goal graph*: for a system of RSS feed reader and a system of reception supporting system for a hospital. The sizes of graphs are shown in Table1. RSS Reader has tags labeled Efficiency, Functionality, Security and Usability, and Reception system has tags labeled Functionality, Maintainability, Reliability, Security and Usability.
- *Target for comparison:* matrix of stakeholders and concerns. Fig.6 is a matrix for the RSS Reader. They are provided as Excel sheets. The columns are ID for goal, goal description, semantic tags, and preferences of stakeholders.
- *Task:* Subjects are asked to answer 17 questions. They can spend two minutes for each and twenty five minutes in total. Twelve questions are expected to be answered in the worst case. Questions are such as follows:

  - ➢ Which concerns is nobody interested in?
  - ➢ Which concerns is only one stakeholder interested in?

➢ Which concerns is everybody interested in?
➢ Which concerns are all developers interested in?

The questions are designed to lead the findings shown in section 4.2.

- *Subjects:* two students, one researcher, and one software engineer. Each subject answers for two goal graphs. The subject does the task with the tool for one graph and with Excel sheet for another graph. They are assigned tasks as Table2.

	A	B	C	D	E	F	G	H	I	J	K	L
1				Dev1	Dev2	Dev3	User1	User2	User3	Admin1	Admin2	Admin3
2	N50	3秒以内に更新される	Efficiency	0	8	0	9	0	0	9	9	0
3	N32	2種類のユーザを想定する	Functionality	0	6	7	3	9	7	0	7	9
4	N9	2重登録はなしにする	Functionality	3	7	5	6	9	0	10	9	9
5	N19	何人が登録しているか見られる	Functionality	0	9	9	2	9	0	2	9	9
6	N4	自分のフィードの閲覧が可能	Functionality	0	8	9	0	7	8	4	8	9
7	N5	登録する	Functionality	0	9	8	0	9	7	8	9	8
8	N22	ネットワーク上のほかのユーザの	Functionality,Security	6	8	7	6	8	9	5	7	9
9	N7	パスワード	Security	0	0	0	2	0	0	8	0	0
10	N25	フィードの公開非公開を設定可	Security	0	0	0	0	0	0	8	0	0
11	N33	ROMユーザは未登録でも利用可	Usability	0	0	8	0	0	9	3	0	0
12	N10	ジャンルわけ可能	Usability	8	0	9	5	0	7	3	0	0
13	N12	タイトル順	Usability	0	0	8	0	0	9	3	0	0
14	N18	ボタンひとつで変更可能	Usability	8	0	8	8	0	9	8	9	9
15	N11	自分で分類	Usability	0	8	8	0	0	9	0	0	0
16	N13	自分の見たいエントリを選んで登	Usability	0	0	9	0	0	3	0	0	0
17	N3	情報の登録はユーザ登録するこ	Usability	0	0	7	7	0	8	0	0	0

FeedReader

Fig. 6. Matrix of stakeholders and concerns

Results. The fourth column in Table2 shows the number of answered questions, the fifth column shows the number of correct answers and the last column shows the time period to finish the task.

Table 2. Task assignment and Results

Subject	Graph	Tool or Table	Completed	Correct	Time
Student1	FeedReader	Tool	17	17	7m10s
	Reception	Table	17	17	8m36s
Student2	FeedReader	Table	17	15	13m37s
	Reception	Tool	17	17	5m58s
Researcher	FeedReader	Tool	17	17	6m19s
	Reception	Table	17	17	17m53s
Engineer	FeedReader	Table	17	17	13m11s
	Reception	Tool	17	17	10m35s

One subject got two wrong answers with table and the others got correct answers for all questions. All subjects have completed the task in short period with the tool than with table. The total time for FeedReader with tool is 13 minutes and 29 seconds and the total time with table is 26m47s. The total time for Reception with tool is 16m33s and the total time with table is 26m29s.

The number of subjects is very small, but the result suggest the tool enables users to identify imbalance of stakeholders or lack of stakeholders faster and more correctly than a matrix of stakeholders and concerns.

5 Related Works

Schneider [18] proposes a technique used to visualize informal communications in a project. Heim [8] proposes a technique how the relationships between requirements are visualized using graph structures. Rohleder [17] proposes a graphical representation on non-functional requirements impact by eight composite rules and seven types of links.

Stakeholder management has a long history [16] but there is no graph representation of relationships between stakeholders and their concerns. Pouloudi and Whitley [15] suggest four principles of stakeholder identification, and describe and approach which is based on these principles, and which they used to identify stakeholders in the drug use management domain. Lyytinen and Hirschheim [13] also provides some guidance for stakeholder identifications for IS, while acknowledge in that the identification of the set of all stakeholders is far from trivial. Sharp at el [19] propose a technique to discovering all relevant stakeholders of a specific system. Our proposing tool will be able to support the Sharp's technique if the model of the technique is applied instead of the semantic tags of AGORA goals. However the Sharp's technique must be done separately in requirements analysis. Our tool is combined with goal graph editor and missing stakeholders will be identified in the middle of requirements analysis.

6 Summary and Future Works

This study presented a technique for visualizing stakeholders concerns in a project using an anchored map. The tool produces a graph which represents the relationship between stakeholder and system concerns. The tool can switch anchors and free nodes, and stakeholder preferences for goals are aggregated with semantic tags. Requirements analysts can easily obtain an overview of the analysis, to find which concern has high priority for end users or system administrators, or to identify the concerns which the smallest number of stakeholders are interested.

References

1. Amyot, D.: Introduction to the user requirements notation: Learning by example. Computer Networks 42(3), 285–301 (2003)
2. Anton, A.I., Potts, C.: The use of goals to surface requirements for evolving systems. In: Proc. International Conference on Software Engineering, pp. 157–166 (1998)
3. Cares, C., Franch, X.: Toward a framework for improving goal-oriented requirement model quality. In: Workshop on Requirements Engineering (2009)
4. Chung, L., Nixon, B.A., Yu, E., Mylopoulos, J.: Non-Functional Requirements in Software Engineering. Springer, Heidelberg (1999)
5. Dardenne, A., van Lamsweerde, A., Fickas, S.: Goal-directed requirements acquisition. In: Science of Computer Programming, pp. 3–50 (1993)
6. Eades, P.: A heuristic for graph drawing. Congressus Numerantium 42, 149–160 (1984)
7. Finkelstein, A., Dowell, J.: A comedy of errors: the london ambulance service case study. In: Proc. 8th International Workshop on Software Specificaiton and Design, pp. 2–4 (1996)

8. Heim, P., Lohmann, S., Lauenroth, K., Ziegler, J.: Graph-based visualization of requirements relationships. In: Proc. Requirements Engineering Visualization, pp. 51–55. IEEE Computer Society, Los Alamitos (2008)
9. ISO/IEC: ISO/IEC 9126-1:2001. Software engineering – Product quality – Part1:Quality model. ISO/IEC (2001)
10. Kaiya, H., Horai, H., Saeki, M.: AGORA: Attributed goal-oriented requirements analysis method. In: Proc. Requirements Engineering Conference, pp. 13–22 (2002)
11. Kaiya, H., Shinbara, D., Kawano, J., Saeki, M.: Improving the detection of requirements discordances among stakeholders. Requirements Engineering 10(4), 289–303 (2005)
12. van Lamsweerde, A.: Goal-oriented requirements engineering: A guided tour. In: Proc. Requirements Engineering Conference, pp. 249–263 (2001)
13. Lyytinen, K., Hirschheim, R.: Information systems failures–a survey and classification of the empirical literature, pp. 257–309. Oxford University Press, Inc., New York (1987)
14. Misue, K.: Anchored maps: Visualization techniques for drawing bipartite graphs. In: Jacko, J.A. (ed.) HCI 2007. LNCS, vol. 4551, pp. 106–114. Springer, Heidelberg (2007)
15. Pouloudi, A., Whitley, E.A.: Stakeholder identification in inter-organizational systems: gaining insights for drug use management systems. European Journal of Information Systems, 1–14 (March 1997)
16. Preble, J.F.: Toward a comprehensive model of stakeholder management. Business and Society Review 110(4), 407–431 (2005)
17. Rohleder, C.: Visualizing the impact of non-functional requirements on variants: A case study. In: Proc. Requirements Engineering Visualization, pp. 11–20 (2008)
18. Schneider, K., Stapel, K., Knauss, E.: Beyond documents: Visualizing informal communication. In: Proc. Requirements Engineering Visualization, pp. 31–40. IEEE Computer Society, Los Alamitos (2008)
19. Sharp, H., Finkelstein, A., Galal, G.: Stakeholder identification in the requirements engineering process. In: Proc. Tenth Int. Database and Expert Systems Applications Workshop, pp. 387–391 (1999)
20. Ugai, T., Aoyama, K.: Organization diagnosis tools based on social network analysis. In: Smith, M.J., Salvendy, G. (eds.) HCI International 2009. LNCS, vol. 5617, pp. 181–189. Springer, Heidelberg (2009)

VICPAM: A Visualization Tool for Examining Interaction Data in Multiple Display Environments

Roshanak Zilouchian Moghaddam and Brian Bailey

University of Illinois at Urbana-Champaign,
201 N Goodwin Ave. Urbana, IL, USA 61801
{rzilouc2,bpbailey}@illinois.edu

Abstract. Multiple Display Environments (MDEs) facilitate collaborative activities that involve the use of electronic task artifacts. Supporting interactions and infrastructures have matured in recent years, allowing researchers to now study how the use of MDEs impacts group work in controlled and authentic settings. This has created a need for tools to understand and make sense of the resulting interaction data. To address this need, we have designed and developed a new interactive analysis tool called VICPAM. Our tool reduces the effort necessary to analyze and make sense of users' interaction data in MDEs. VICPAM consists of several components: (i) a *time-aligned view*, which shows users' activities over time and the duration of each activity; (ii) A *spatial view*, which gives a 2D overview of all users' activities in the environment; (iii) A *time-bar*, which allows selection of a desired time period for in-depth analysis; and (iv) a *video player,* which allows the user to watch a video of the session synchronized with the selected period of time.

Keywords: Visualization, Multiple Display Environment, Interactive Analysis.

1 Introduction

Multiple Display Environments (MDEs) represent an emerging environment for conducting group problem solving activities involving electronic task artifacts. Group-based programming and collaborative design are two examples. During these activities, group members will relocate cursor input and artifacts such as IDEs, web pages, sketches, and documents between personal and shared devices. This provides the opportunity to share, manipulate, and discuss related information. In MDEs each individual typically has a personal device, and all computers are connected via a network [11]. Large shared displays may also be present to facilitate group collaboration. Colab [11], iRoom [9], and OCEAN-Lab [12] are examples of such environments. Figure 1 shows an example of a MDE configured for collaborative software development.

With many instances of MDEs available, researchers are beginning to study how their use affects collaborative problem solving. For example, Biehl et al. [3] studied the use of MDEs in the context of collaborative software development activities. In another study Streitz et al. compared multiple configurations of MDEs for group meeting tasks [12]. Also Izadi et al studied a system called “Dynamo” which enables

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 278–287, 2011.

users to share and exchange information in unfamiliar public places [8]. In all these studies, in order to evaluate the environment, all interactions between group members and devices were captured, e.g. which applications were shared, what devices were involved, and at what time. The researchers had to analyze the interaction data collected to answer their respective study questions. However, analyzing this type of data can be challenging because multiple interactions of interest can overlap in time (e.g., two or more users placing artifacts on a shared display) and users (e.g., one user sharing an artifact with others at different times). This data can be complex and voluminous.

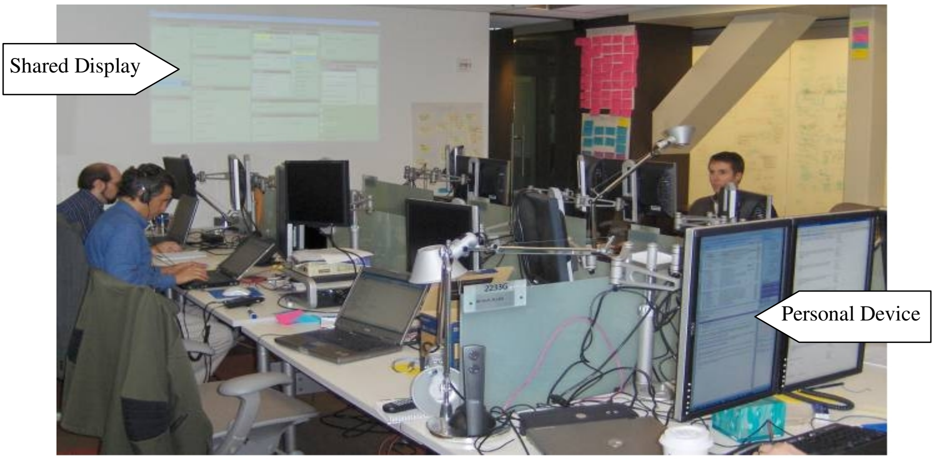


Fig. 1. A group of five developers are using a MDE for collaborative software development (from [4] with authors' permission)

A common method for analyzing interaction data from MDEs (and other) evaluations is the use of spreadsheets. Spreadsheets can be used to record and examine the data in a structured manner, but they do not allow interaction with the resulting visualizations. Visual explanations are suitable to get an overview of the data, but they do not (by themselves) support interaction for investigating patterns and clusters. Interaction techniques can allow the user to better perceive the information when visually exploring a data set [6]. Many interactive visualization tools have been developed to make sense of complex data in myriad domains, including Google Finance [1] for analyzing stocks, profilers and debuggers in software development, and Nasa's Hurricane analysis tool in the domain of earth science [2].

COPROT was the first interactive analysis tool designed for exploring data gathered from studies in MDEs [12]. This tool helps researchers to record and track cooperation modes and the time spent in each mode by watching the video of an evaluation session. Unlike their tool, our work aims to support researchers in determining the cooperation modes and the *collaboration patterns* by both visually representing the interaction data and showing the video of the evaluation session. WorkspaceNavigator [7] enables capture, recall, and reuse of the digital information, decisions, and rationale generated by a group collaborating within an interactive workspace. Our work is original in that it targets exploring data in order to investigate the collaboration and communication patterns between group members.

To address the need to make sense of the interaction data from MDEs, we have designed and developed a new interactive analysis tool called VICPAM (Visualizing Interaction and Communication PAtterns in MDE's). VICPAM consists of (i) a time-aligned view, which shows users' activities and the duration of each activity; (ii) a spatial view, which gives a 2D overview of all users' activities; (iii) a time-bar, which can be used to adjust the period of time under analysis; and (iv) a video of the session synchronized with the timeline. The tool reads an XML file as input.

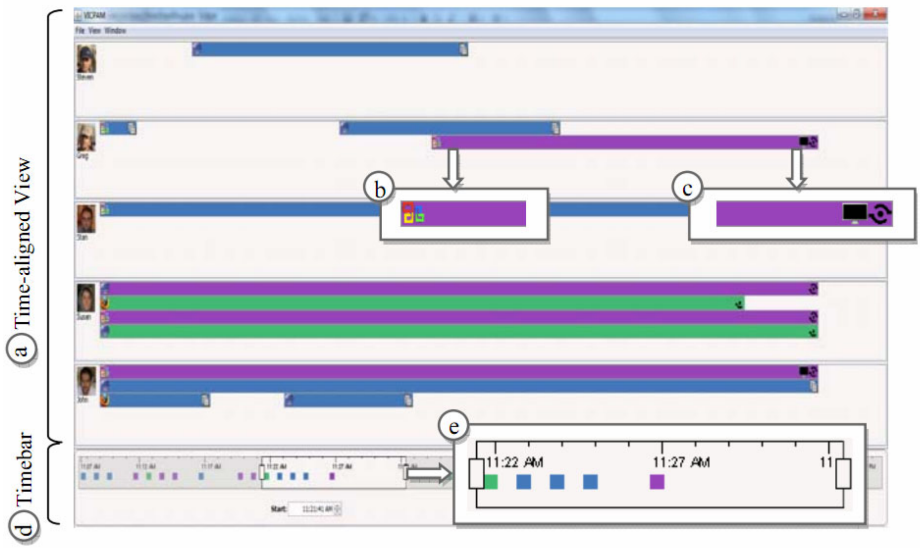


Fig. 2. The main screen consists of the time-aligned view and an interactive visual time-bar. (a) The time-aligned view shows users' activities and the duration of each activity. User's activities are shown by rectangles. (b) The application icon is located to the right of a rectangle while (c) the activity icon is located on the left. When the activity has been performed via a shared display, a device icon is also shown on the left indicating the use of the shared display. (d) The time-bar can be used to select a desired time period for analysis. (e) The square-shaped visual indicators displayed on the time-bar mark the start of a new activity.

The contribution of our tool is that it demonstrates a new interactive visualization technique for facilitating analysis of complex data collected from studies of group problem solving in MDEs.

2 VICPAM

VICPAM is an interactive visualization tool which helps researchers to analyze the interaction data captured during their studies in MDEs. Our system was developed through an iterative design process, starting with examining different configurations of MDEs and several studies conducted in each configuration. We designed two different prototypes that addressed the requirements identified in our analysis. To evaluate them, two MDE researchers and three HCI researchers were recruited, asked to perform similar tasks (e.g., identify the duration of a specific activity) with each prototype, and asked to explain their relative strengths and weaknesses. From the results, we implemented our prototype of VICPAM.

VICPAM reads the interaction data from an XML file. The file begins with meta-data information including the number of users and their information, their relative position in the room, types of activities users can perform in the environment, and a link to the icon of each activity type to be displayed in the visualization. The actual data should consist of a <total duration> tag which indicates the duration of the entire session as well as a number of <entry> blocks which represent an activity a user has performed on a specific application for a particular period of time. In each <entry> block there should be a <user> tag indicating the user's name, an <application> tag indicating the application's name, an <activity> tag indicating the type of activity the user has performed, an optional <device> tag indicating the device used for performing the activity (personal device or shared display), and a <time> tag which contains <start> and <end> tags specifying the duration of the activity.

2.1 Data Background

We will describe VICPAM in context of the interaction data collected in studying IMPROMPTU, an interaction framework for MDES, for collaborative software development [4]. The data contains three different user activities: (i) *Show* which means an application window is provided to the group in view-only mode. (ii) *Share* which means an application window is provided to the group members and anyone can edit its content if they replicate it on their own machine; (iii) *Replication* which means an instance of a shared or shown application has been copied on a personal machine. IMPROMPTU logs the usage data including the type of the activity (i.e. share, show, and replication) and the duration of the activity. All of the figures illustrating the use of our system are based on this data set.

2.2 Visualization

VICPAM's main screen consists of the time-aligned view and an interactive visual time-bar (Figure 2). The time-aligned view shows the exact duration of each activity a person has performed (Figure 2a). Each person is displayed inside a rectangular frame. This frame includes the person's information, and a set of rectangles representing the activities he performed on different applications. The width of each rectangle corresponds to the duration of the activity and the color corresponds to the type of activity. For this particular data set, green represents show, purple represents share, and blue represents replication. There are also a number of icons shown on each rectangle including the application icon on the right side (Figure 2b) and the activity icon on the left (Figure 2c). If the activity has been performed via the shared display, a device icon will also be shown on the left side indicating the use of the shared display. Note that the color and icons used for each activity type are specified at the beginning of the input XML file to make the tool usable for different data sets.

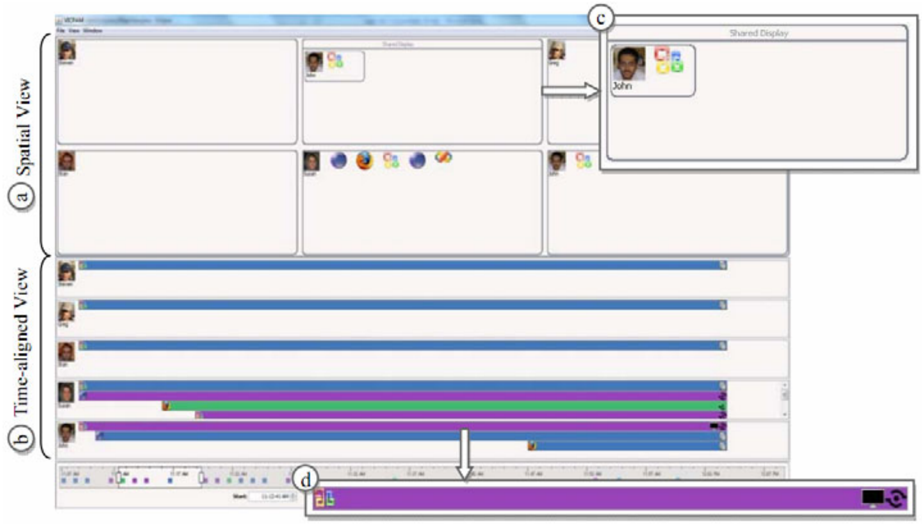


Fig. 3. When the user selects "Show Spatial View", the spatial view will be displayed on the top of the time-aligned view. (a) The spatial view shows a spatial 2D view of the users and their activities. (b) The time aligned view shows the exact duration of activities. (c) As it's shown in the shared display section of the spatial view, John has shared an instance of visual studio in the shared display. (d) John's activity is also shown in the time-aligned view.

The user can use the slider in the time-bar to select a desired period, the time period can be as small as a second or as large as the entire session (Figure 2d). Upon her selection all the views in the visualization will be updated illustrating only the activities happened during the selected period. To support the user in detecting and exploring interesting periods, the time-bar includes visual indicators which mark the start of a new activity (Figure 2e).

By selecting "Show spatial view" from the "View" menu on the menu-bar, a spatial view is added to the visualization (Figure 3). The spatial view shows a 2D overview of the activities people have performed in the environment. Each person is displayed in a rectangular frame, which includes a picture of the person, her name, and the applications that she has shared or shown during a specific period of time. Unlike the time-aligned view, the spatial view doesn't show the duration of the activities, instead it orders the people based on their actual position in the room. For example, this can be used to determine whether a person's position affects his or her collaboration patterns. Also, the spatial view is more useful in answering certain questions, such as what is the most frequently used application in the entire session. Both the spatial view and the time-aligned view are adjustable, these views can be opened in new windows, and the size and place of the artifacts inside these views can be configured.

To help the researcher with data analysis, several interaction mechanisms have been implemented in the interface:

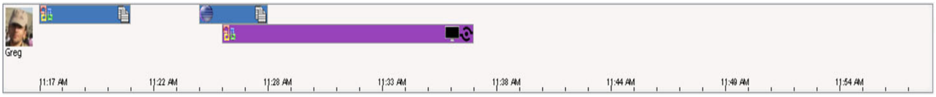


Fig. 4. Selecting "Show rulers" from the "View" menu in the menu-bar displays a time ruler under each person's frame

Interaction with users and activities: Hovering over each user's picture shows detailed information about that user, including the number of activities she has performed. Hovering over each activity shows information about that activity including the exact start and end time, the application name, and the user's name.

Display time rulers: Selecting "Show rulers" from the "View" menu in the menu-bar displays a time ruler under each person's frame in the time-aligned view (Figure 4). Using these time rulers the researcher can easily compare the start and end time of activities without having to draw an imaginary vertical line down to the time-bar.

Link between the spatial view and the time-aligned view: To support the researchers in using both spatial view and time-aligned view simultaneously for analysis, a virtual link has been created between the users and activities in these views. When both views are open, hovering over a user's picture in the time-aligned view will highlight that user and his activities in the spatial view. Conversely, hovering over a user's picture in the spatial view will highlight that user and his activities in the time-aligned view. Hovering over an activity on the time-aligned view will highlight the corresponding activity in the spatial view and vice versa.

2.3 Video of Interaction

A MDE's evaluation session is usually captured by a video camera. This video can be imported to VICPAM and played in a separate window. The video is synchronized with the time-bar. When the user presses the play button the part of the video corresponding to the selected time period in the time-bar begins to play.

2.4 Filtering

A separate filtering window is available to filter the data based on applications, users, and activities (Figure 5). For instance, if the researcher selects one application and filters the rest, only the activities related to the selected application will be shown in the interface. The filtering options allow researchers to focus on their desired part of the data. This feature is especially helpful when the amount of data shown in the interface is overwhelming. During our preliminary evaluations, the participants found the filtering options to be helpful for answering focused questions (e.g. how many applications have been shared by a particular user).

2.5 Putting It Altogether

In this section, we illustrate the value of our tool through a user scenario. The scenario is grounded in the same data set described earlier in Data Background.

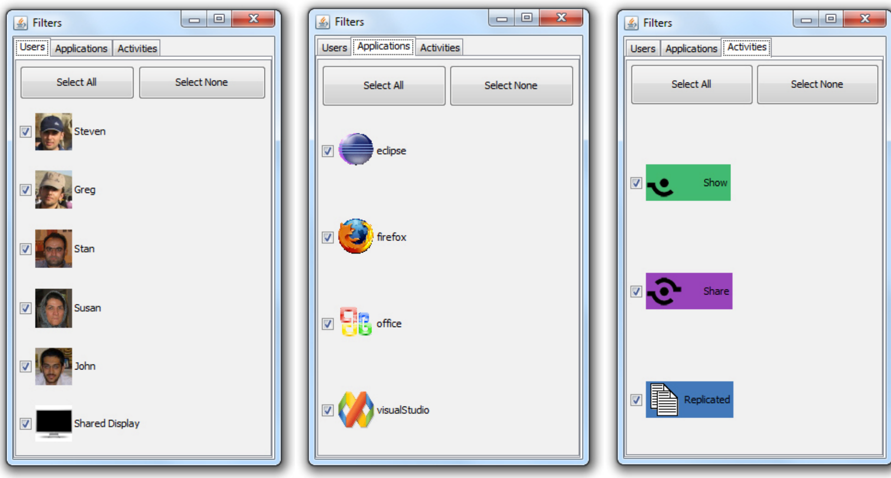


Fig. 5. The filtering window has three tabs for filtering out users, applications, and activities

Bob is a researcher interested in understanding how the use of an MDE affects collaborative software development. Of the many interaction frameworks available, Bob installs IMPROMPTU [4] to enable his study of MDEs. In his study, five developers utilize the MDE for four hours each day for two weeks. The MDE consists of five personal devices and a large shared display positioned such that all the developers can see it. During each evaluation session, he logs all of their activities and captures a video of each session. Now he wants to analyze the data collected to examine the use of the MDE. He writes an adaptor to convert each session's collected data to the XML format readable by VICPAM, He opens VICPAM and loads the first evaluation session by clicking on "File" and then "Load Session" on the menu-bar.

First Bob wants to determine which application is most frequently used in this session. He selects the entire session by using the slider in the time-bar. Then, he opens the spatial view and counts the number of times each application has been displayed in this view. He discovers that visual studio was the most frequently used application during the session. Then he would like to find the most active user. Again, using the spatial view while the entire session is selected, he identifies the most active user by counting the number of application icons in front of each user's representation. He learns that Susan was the most active user in the session.

Now he would like to examine the collaboration patterns among developers, e.g. whether all team members work together on a particular problem or there are two or more smaller groups of people who work separately on different aspects of the problem. He closes the spatial view and focuses on the time-aligned view. He frequently changes the selected time period using the slider and inspects the developer's activities. He detects an interesting pattern in the middle of the session. He sees in the time-aligned view that John, one of the developers, had shared a Microsoft Office application on the shared display and all team members replicated a copy of that on their personal machines. To explore this further, Bob plays the synchronized video and sees that the developers took turns to add the features for the next iteration of their software. This activity was marked as an example of group collaboration.

He continues changing the selected time period until he reaches the end of the session. He finds another interesting pattern near the end where Steven, a developer, has shared an instance of his Visual Studio and Greg has replicated that shared application on his own machine. By playing the video of their interaction, Bob discovers that they are reviewing Steven's code. Bob notes this activity as an example of peer-to-peer collaboration.

Using the views and interaction mechanisms provided in VICPAM, Bob was able to quickly answer his questions about the most frequently used software and the most active developer. He was also able to identify two interesting collaborative patterns that he can report. This was possible because VICPAM was designed to analyze the data captured from MDE evaluations. With other analysis techniques, Bob would not have been able to answer his research questions as easily.

3 Discussion

VICPAM is fully implemented in Java and consists of about 2,500 LOC. The user interface was developed in Java Swing and the video player was implemented using Java Media Framework (JMF). In order to make the code extensible, it was written in three layers; data, domain, and user interface (or presentation). The layered architecture enables other researchers to add new features or extend the tool to handle different data sets.

Three major iterations of development and evaluation were performed on VICPAM. After each evaluation, we modified the interface according to the lessons learned. For example, during one evaluation, we learned that users prefer to initiate their analysis from the time-aligned view rather than the spatial view. We therefore implemented the time-aligned view as the default and provided an option of adding the spatial view via a menu. We also implemented the time rulers to further support this preference.

Our tool was evaluated with only one set of data from a particular framework, IMPROMPTU [4]. However, IMPROMPTU is a highly flexible framework as it allows any existing application to be utilized in an MDE and also works with any combination of personal and shared devices. Therefore this framework can be used in other studies in MDEs. Also, the data collected from other MDE frameworks can be converted to the format compatible with VICPAM. Either the framework needs to be instrumented to log the interaction data in a compatible format, or an adaptor needs to be written that converts the existing format into a compatible format. VICPAM can even be adapted for analyzing data collected from different user and hardware configurations. For example, the tool can be used to analyze data collected from studies of single display groupware such as tabletops. There have been many evaluations of interactive tabletops for collaborative work and play [5, 10]. The data collected from these studies can be imported into VICPAM by modeling personal areas as personal devices and modeling the shared working area as a shared display.

4 Conclusions and Future Work

Researchers are beginning to study how the use of multiple display environments affects collaborative problem solving. To foster analysis and exploration of data gathered from these studies, we have designed and developed an interactive analysis tool called VICPAM.

VICPAM allows researchers to investigate different aspects of the interaction data by providing two different views of the data set: the time-aligned view and the spatial view. The time-aligned view shows people's activities and the exact duration of the activity. This view helps researchers detect collaboration patterns between users. The spatial view displays the users based on their positions in the workspace. This view provides an overview of the activities performed in the selected period of time. It can also be used to determine if users' positions affect their communication and collaboration patterns. Besides these views, a separate filtering window enables the researchers to filter unnecessary information and focus on their desired part of the data.

Although we built and tested our tool with a data set generated from the use of one interaction framework in one task domain, results from our study did reveal opportunities for improving the design of this type of interactive analysis tool. One improvement would be to add another view that includes statistical data either in the form of graphs or spreadsheets. For example a graph showing the total amount of activities at any point in time can help to identify busy periods as potentially interesting periods to investigate. Another improvement would be to provide data converters that could map the data gathered from studies conducted using different interaction frameworks and device configurations to the data format that our tool accepts. By using speech recognition and natural language processing techniques, group member statements related to their activities could be extracted from the session video and associated with their activity in the visualization. Finally, if the data set also includes screenshots of the actual documents that group members are sharing, each activity in the visualization could provide a link to the actual document shared.

References

1. Google finance, `http://finance.google.com`
2. Nasa Hurricane data analysis tool, `http://disc.sci.gsfc.nasa.gov/daac-bin/hurricane_data_analysis_tool.pl`
3. Biehl, J.T., Bailey, B.P.: Improving interfaces for managing applications in multiple-device environments. In: Proceedings of the Working Conference on Advanced Visual Interfaces, pp. 35–42. ACM, New York (2006)
4. Biehl, J.T., et al.: Impromptu: a new interaction framework for supporting collaboration in multiple display environments and its field evaluation for co-located software development. In: Proceeding of the Twenty-Sixth Annual SIGCHI Conference on Human Factors in Computing Systems, pp. 939–948. ACM, New York (2008)
5. Fiebrink, R., Morris, D., Morris, M.R.: Dynamic mapping of physical controls for tabletop groupware. In: Proceedings of the 27th International Conference on Human Factors in Computing Systems, pp. 471–480. ACM, New York (2009)
6. Hibbard, W., et al.: Interaction in Perceptually-Based Visualization. In: Grinstein, G.G., Levkowitz, H. (eds.) Perceptual Issues in Visualization, pp. 23–32 (1995)

7. Ionescu, A., Stone, M., Winograd, T.: Workspace Navigator: Tools for Capture, Recall and Reuse using Spatial Cues. In: Interactive Workspace 2002, Stanford University (2002)
8. Izadi, S., et al.: Dynamo: a public interactive surface supporting the cooperative sharing and exchange of media. In: Proceedings of the 16th Annual ACM Symposium on User Interface Software and Technology, pp. 159–168. ACM, New York (2003)
9. Johanson, B., Fox, A., Winograd, T.: The Interactive Workspaces Project: Experiences with Ubiquitous Computing Rooms. IEEE Pervasive Computing, 2002 1(2), 67–74 (2002)
10. Morris, M.R., Lombardo, J., Wigdor, D.: WeSearch: supporting collaborative search and sensemaking on a tabletop display. In: Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work, pp. 401–410. ACM, New York (2010)
11. Stefik, M., et al.: Beyond the chalkboard: computer support for collaboration and problem solving in meetings. Commun. ACM 30(1), 32–47 (1987)
12. Streitz, N.A., Rexroth, P., Holmer, T.: Does "roomware" matter?: investigating the role of personal and public information devices and their combination in meeting room collaboration. In: Proceedings of the Fifth Conference on European Conference on Computer-Supported Cooperative Work, pp. 297–312. Kluwer Academic Publishers, Dordrecht (1997)

Part IV

Security and Privacy

Privacy Concern in Ubiquitous Society and Research on Consumer Behavior

Yumi Asahi

Department of Management Science, Faculty of Engineering,
Tokyo University of Science
1-3,Kagura-zaka, Shinjuku-ku, Tokyo, 162-8601
asahi@ms.kagu.tus.ac.jp

Abstract. Ubiquitous means being present everywhere at the same time, which expresses the situation that users can have no trouble accessing computers and networks from anywhere, anytime. This broadens its prospect in marketing. On top of the customer information and buying history that the company keeps track of, the customers' environment information including time, place and activity field has become available. These days, due to ubiquitous marketing, concierge-styled service becomes a real possibility. This is where various suggestions are provided according to users' interests, thoughts and behavioral patterns. This research explains consumer awareness based on a questionnaire survey about consumers' privacy in ubiquitous marketing.

Keywords: Classifying respondents, Cluster analysis, Covariance structure analysis.

1 Introduction

Anyone can access networks to obtain information and services no matter when they need them. This means various information can be collected by anyone, anywhere, any time. In other words, personal information associated with privacy is also collected [2]. Consumer information such as age, address, sex and annual income is utilized in existing one-to-one marketing [1]. Ubiquitous society allows us to collect and analyze a large amount of consumers' personal information for marketing [4]. The information includes real time location, buying history, status of product use, etc. Although the Private Information Protection Law was enacted in 2005, it is still difficult to obtain consent for the use of personal information. The situation of unsolicited email is not yet resolved [3]. What information do people hesitate to provide and what information are they willing to provide in the name of privacy? This research explains consumer awareness based on a questionnaire survey about consumers' privacy in ubiquitous marketing.

2 Questionnaire Survey on Privacy

In this research, a questionnaire survey was conducted with students who have previously used ubiquitous marketing services. In this survey the target was 174 aged

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 291–300, 2011.

19 or over living in the Tokyo metropolitan area. Survey sheets given to the targets by hand were collected after filling out the answers. It was comprised of 25 questions about privacy awareness. The questions were related to 'status of ubiquitous availability', 'providing privacy information' and 'privacy information leaks'. Major items were answered using five stage assessment system except for the question regarding respondent information. Respondents to the data to be used in the analysis are 174 persons (male: 106, female: 68). The following items are included in the question items:

Q1. “Personal data may be used only if it provides any benefit as long as it isn’t misused.”
Q2. “Personal data may be disclosed depending on a condition as long as the individual isn’t identified.”
Q3. “To what extent do you agree to disclose your information to receive data distribution from stores?”
Q4. “To what kind of stores do you agree to provide information?”
Q5. “With what kind of merchandise information do you agree to provide?”
Q6. “How do you feel about information provision based on personal data utilization which is useful for you?”
Q7. “Do you want to use such system as to distribute information from stores for you to select a store when you eat out?”
Q8. “Do you want to use such system as to distribute information from stores for you to select a store when you go shopping?”
Q9. “Do you want to use such system as to distribute information from stores for you when you go to supermarkets?”
Q10. “Have you ever used any such information provision service as described from Q7 to Q9?”
Q11. “I feel my personal data is used for other purpose while I’m unaware.”
Q12. “How do you tolerate for leakage of your registration information?”
Q13. “Does what kind of privilege make you agree to disclose your personal data?”
Q14. “Assuming a situation where you’re shopping at a supermarket, how frequently do you want to receive discount information?”
Q15. “While you aren’t shopping, how frequently do you want to receive discount information?”

3 Analysis on Privacy

3.1 Flow of Analysis

First, the author demonstrates a relationship between willingness to use distributed information and views on personal data among respondents using covariance structure analysis. Then, the author analyze whether any difference is found in answer items for the questions by classifying respondents based on their views on personal data using cluster analysis. Further, the author analyze whether any difference is found in answer items by gender.

3.2 Relationship between Views on Personal Data and Willingness to Use Information Distribution Service

As a certain level of strong correlation has been found between four variables, i.e. Q6 Information service based on personal data utilization, Q7 Whether you want to use information distribution service from stores in selecting a store when you eat out, Q8.Whether you want to use information distribution service from stores in selecting a store in case of shopping and Q9 Whether you want to use information distribution service from stores when you go to supermarkets, a model is shown in Fig.1, which depicts how consciousness to personal data has impacts on “motivation for information distribution service utilization”, based on an assumption that a potential consciousness as the “motivation for information distribution service utilization” exists behind these four variables.

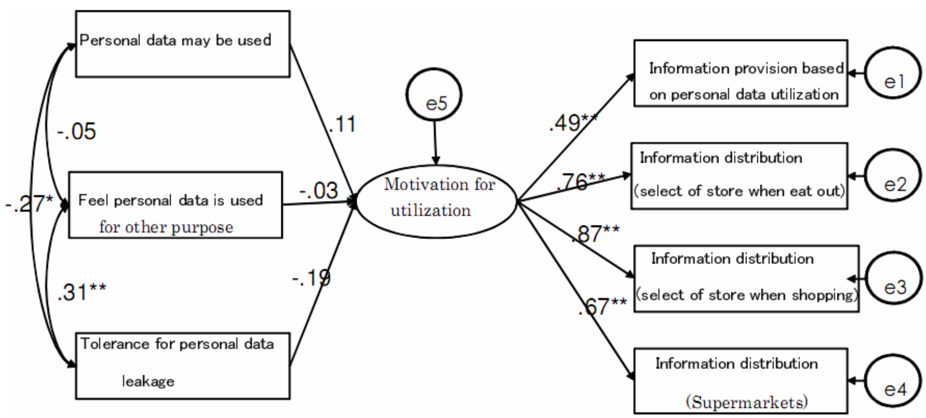


Fig. 1. Model of Casual Relation Between Consciousness to Personal Data and Motivation for Information Distribution Utilization

In the Fig.1, figures with ** mark and *mark indicate that the path coefficients are significant at the 1% level and 5% level respectively and those without any mark indicate that it isn’t significant at the 5% level. Since p-value is at isn’t rejectable level of 0.785 based on fit index, the model is acceptable. In addition, GFI and AGFI were 0.977 and 0.940 respectively. The model is determined to be well applicable, since a model is supposed to be better applicable as both GFI and AGFI get closer to 1[5] [6].

In terms of path coefficient, any path to “motivation for information distribution service utilization” from each variables, i.e. “personal data may be used”, “I feel personal data is used for other purpose” and “generosity for personal data leakage”, wasn’t significant. Therefore, it is conceivable that those who agree that personal data may be utilized don’t necessarily want to receive information provision based on personal data, and that those who feel it unforgivable that personal data is leaked don’t necessarily regard information service based on personal data to be needless. In other words, it may be said that there isn’t strong relationship between “view on personal data” and “motivation for utilization of information service based on personal data”.

Then, we perform simultaneous analysis of multiple populations by dividing population into two groups by gender using the model in Fig.1. Judging from the fit indexes resulted from analysis result, such as p-value of 0.441, GFI of 0.926 and AGFI of 0.833, the model is determined to be well applicable. As a result of the analysis, any path to "motivation for information distribution utilization" from "personal data may be used", "I feel personal data is used for other purpose" and "generosity for personal data leakage" wasn't significant at the 5% level neither in male population nor in female population. It may indicate that there is no relationship between "view on personal data" and "motivation for utilization of information service based on personal data". Correlation coefficients between "personal data may be used", "I feel personal data is used for other purpose" and "generosity for personal data leakage" are shown for each population in Table 1.

Table 1. Path coefficient in Multiple Populations

Feel is is used for other purpose	⇔	For personal data leakage	0.415**	0.119
Personal data may be utilized	⇔	Feel it is used for other puproses	-0.259	-0.273
Personal data may be utilized	⇔	Tolerance for personal data leakage	-0.117	0.385*

**: significant at 1%, *:significant at 5%. :significant at 10%

Positive correlation is found between "I feel personal data is used for other purpose" and "generosity for personal data leakage" in male population, whereas significant correlation isn't recognized in female population. It may indicate that "those who feel it unforgivable to leak personal data" have a stronger tendency to think as "I feel personal data is used for other purpose" in case of men. In case of women, on the other hand, it may be said that "those who feel it unforgivable to leak personal data" have a stronger tendency to think as "I don't want personal data to be used", since there is a positive correlation between "personal data may be used" and "generosity for personal data leakage" in case of women while there isn't significant correlation between them.

3.3 Feature Analysis by Group

Classification of Respondent. Subsequently, the author perform a cluster analysis for classifying respondents by their views on personal data. Using three variables, i.e. "personal data may be used (five-grade evaluation)", "I feel personal data is used for other purpose (five-grade evaluation)" and "generosity for personal data leakage (five-grade evaluation)", Ward method has been used to combine clusters. As a result of the analysis, respondents have been classified into three groups. Mean value of each group is shown in Fig.2, and features interpreted by number of respondents and name of each group as well as by mean values are shown in Table 2.

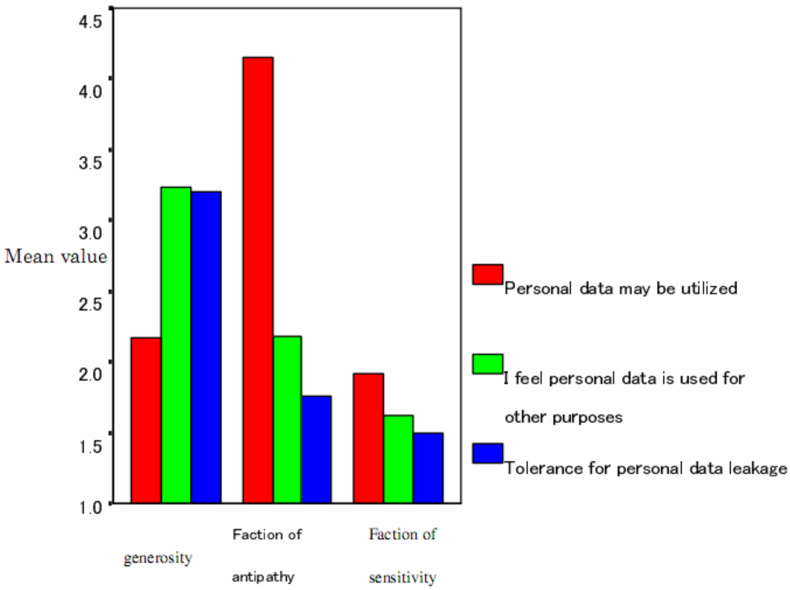


Fig. 2. Mean values of variables in each group

Table 2. Characteristics in each group

	Number of members	Name of group	Features
Group 1	60	Faction of generosity	Those who agree with exploitation of personal data and don't care so much about usage of personal data for other purpose as well as its leakage
Group 2	66	Faction of antipathy toward usage	Those who neither accept personal data to be exploited nor feel its leakage to be forgivable
Group 3	48	Faction of sensitivity	Those who accept personal data to be exploited but wonder if it is used for other purpose and feel its leakage to others to be unacceptable

Feature Analysis on Groups. Subsequently, the author analyze what find of features are recognized in each group. Cross-tabulation about group and gender is shown in Table 3. Even though it is recognized in Table 3 that there is a tendency that slightly larger number of men belong to the faction of e and that slightly larger number of women belong to the faction of the faction of antipathy toward usage, it is conceivable that there isn't so statistically significant relationship between gender and group since the result of chi-square test shows each value of $\chi2$ and p to be 1.005 and 0.605 respectively, which means it isn't at a level to be rejected. Further, comparing mean values obtained from t-test which has been performed based on gender for variables of Q1, Q10 and Q11, all of them hadn't significant difference at the 5% significance level.

From this result, it may be said there isn't great difference in consciousness by gender over how personal data is used. Cross-tabulation about group and application experience of information distribution service is shown in Table 4.

Table 3. Relationship between groups and gender

	Group			Total
	Faction of generosity	Faction of antipathy toward usage	Faction of sensitivity	
Gender male frequency Percentage within group	40 66.7%	36 54.5%	30 62.5%	106 60.9%
Gender female frequency Percentage within group	20 33.3%	30 45.5%	18 37.5%	68 39.1%
Total frequency Percentage within group	60 100.0%	66 100.0%	48 100.0%	174 100.0%

Table 4. Relationship between variables and service usage experience

	Group			Total
	Faction of generosity	Faction of antipathy toward usage	Faction of sensitivity	
Usage of information distribution service - experienced frequency Percentage within group	28 46.7%	28 42.4%	16 33.3%	72 41.4%
Usage of information distribution service – not yet experienced frequency Percentage within group	12 20.0%	28 42.4%	22 45.8%	62 35.6%
Usage of information distribution service – unknown frequency Percentage within group	20 33.3%	10 15.2%	10 20.8%	40 23.0%
Total frequency Percentage within group	60 100.0%	66 100.0%	48 100.0%	174 100.0%

In Table 4, number of those who have less application experience tends to be rather smaller in faction of generosity and it is conceivable that there is no relationship since the result of chi-square test shows values of $\chi 2$ and p to be 6.099 and 0.192 respectively, which means it isn't at a level to be rejected. Therefore, it may be said that neither remarkably larger is the number of those who have experienced the use of "information distribution service in which some information is timely sent from stores" nor smaller is the number in certain group. Thus, those who are generous for personal data exploitation don't necessarily use information provision service and those who don't want their personal data to be utilized don't necessarily refuse to use such service.

Further, one-way analysis of variance, which was performed to examine whether there is any difference between mean values of variables of "frequency of service information reception (when shopping)"and "frequency of service information reception (when not shopping)", has proved that mean values aren't different between groups since both values weren't significant at the 5% level. Therefore, it may be said that there isn't much relationship between views on personal data and frequency to receive service information.

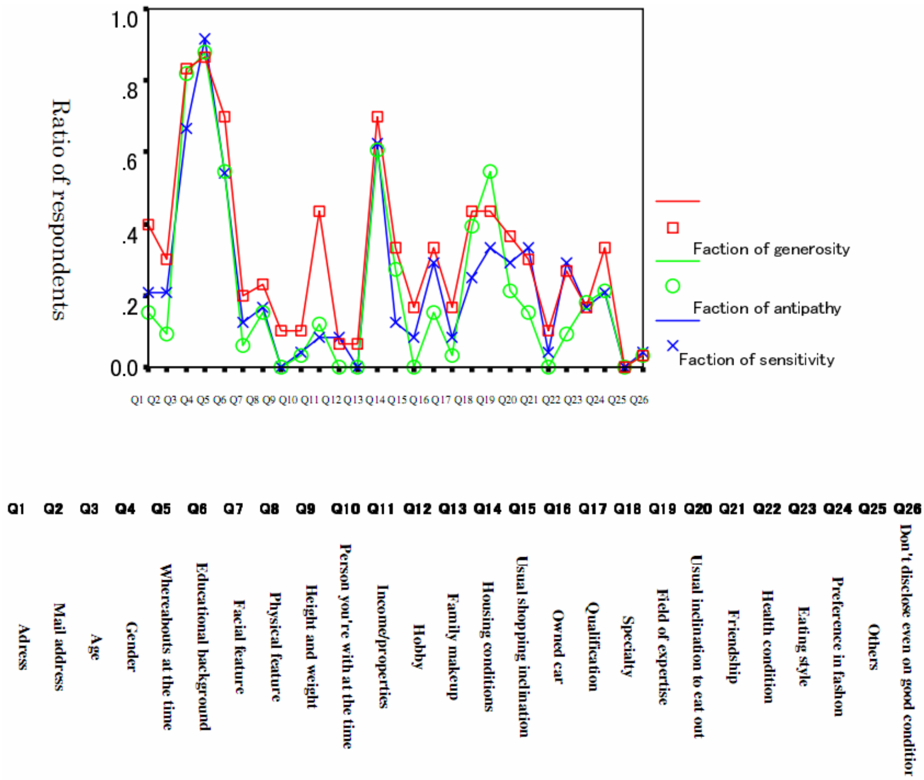


Fig. 3. Ratios of data for which respondens are acceptable to disclose

In terms of a question "what you think it acceptable to be disclosed", ratios of those who answered positively to the question are shown by items in Fig.3. Vertical axis indicates how many respondents answered as "acceptable" when the total is taken as 1.0.

As a whole, it may be recognized that the ratios of those who responded it acceptable to provide information of "age", "gender", "hometown" and "hobby" tended to be higher, whereas that of those who responded positively to "current whereabouts", "features of face and body", "income and asset", "owned car" and "friendship" tended to be lower. In addition, as a tendency by groups, it is recognized that the ratio of those who answered it acceptable that "address" and "mail address" are disclosed is rather higher in "faction of generosity". Moreover, it is also recognizable that even "faction of sensitivity" and "faction of antipathy toward utilization" tended to show higher rate of those who responded it acceptable to provide information of "age", "gender" and " hobby" .The ratio of those who answered acceptable for "information acceptable for providing to stores for their information distribution" is shown by items in Fig.4. As a whole, it is recognizable that the ratios of those who answered it acceptable to provide to the extent of "favorite food" and "visiting frequency/usage history" are higher and the ratios of those who did it acceptable to provide information of "address and name" "annual income and allowance", "whereabouts" and "person you're with" are lower. As a tendency in groups, the ratio of those who agree to provide address and name tends to be rather higher in "faction of generosity", while the ratio of those who responded it depending on benefits such as coupon. However, it is observed as a whole that there isn't such large difference in tendency between groups.

The ratios of those who responded positively to the item of "to what kind of stores do you acceptable to provide such personal data as shown in Q2" are shown by items in Fig. 5. As a whole, the ratio of those who responded it acceptable to provide for "department store", "electronics retail stores" and "restaurants" tended to be higher while that of those who responded it acceptable to provide to "drug stores" lower. As a tendency by groups, the ration of those who responded it acceptable to provide to "electronics retail stores" and "convenience stores" is rather lower in "faction of antipathy toward exploitation" than in other two groups and even the former faction tends to be acceptable to provide their information only to "department stores" and "restaurants".

The ratios of those who responded positively to the item of "about what kind of product information do you accept to provide". As a whole, it is recognizable that those who responded it acceptable to provide information about"foods" tended to be higher, whereas that of those who responded it acceptable to provide information of "livingware" and "medical goods" to be lower. As a tendency by groups, although the ratio of those who responded it acceptable to provide information about only "clothes" tended to be rather higher in "faction of generosity" than in other two groups, it is recognizable that there isn't such large difference in tendency between groups.

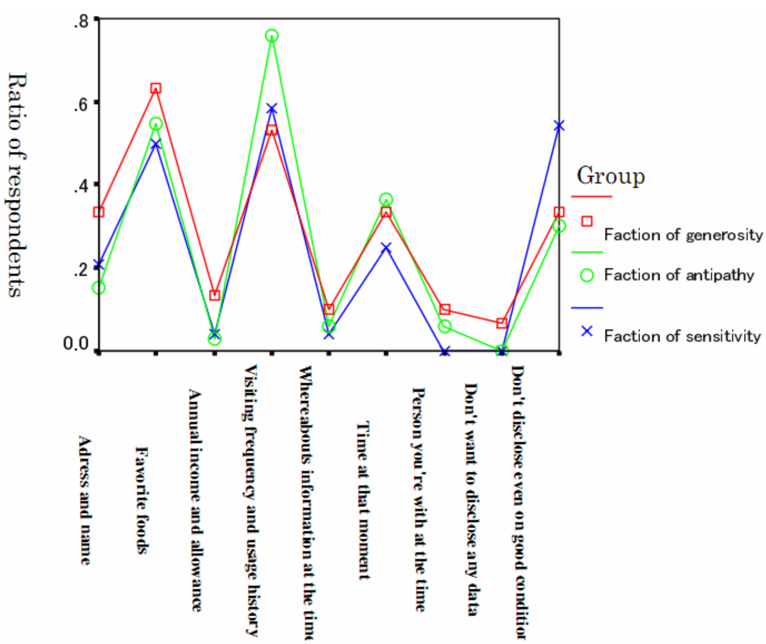


Fig. 4. Ratios of data of respondents which is need for information distribution

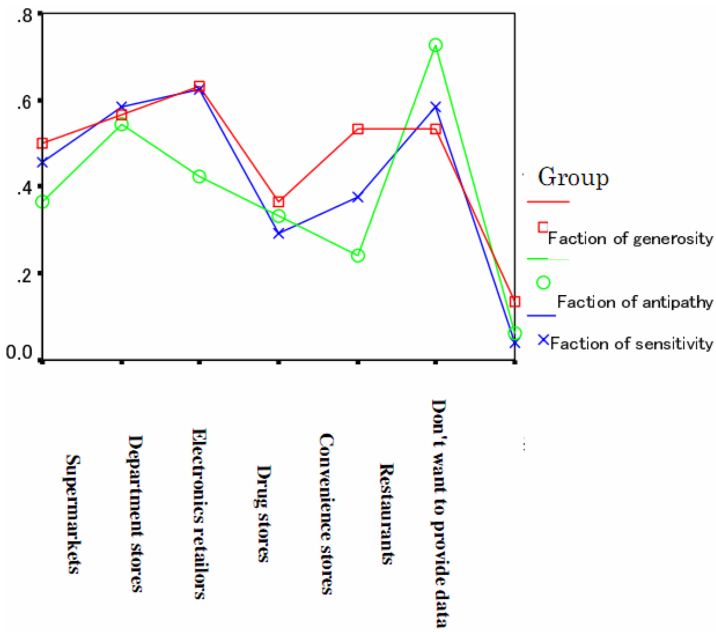


Fig. 5. Ratios of stores respondents are acceptable to provide its information

4 Conclusion

Respondents have been classified into three groups by their views on personal data. Each group has such characteristics that "faction of generosity" which is generous to leakage of personal data to some extent with generosity for its use, "faction of sensitivity" which is unforgivable against leakage of personal data but doesn't mind such data to be used and "faction of antipathy toward utilization" which is unforgivable against leakage of personal data and unacceptable to its use as well.

In terms of information which was thought to be acceptable to be provided to stores for distributing information, it has been proved that more people thought it acceptable to provide "favorite foods" and "visiting frequency and usage history" and that less people thought it acceptable to provide "address and name", "annual income and allowance", "whereabouts information" and "person you're with" as a whole. It has also been proved that "faction of generosity" tended to provide "address and name" more positively than other two groups. In terms of stores for which personal data was allowed to be disclosed, it has been proved that more people thought it acceptable to disclose only "department stores" and "restaurants" and that even "faction of antipathy toward exploitation" tended to think it acceptable to disclose the same as a whole. In terms of product information which was acceptable to be provided, it has been proved that more people tended to think it acceptable to provide only "foods" and that less people tended to think it acceptable to provide "livingware" and "medical goods" as a whole. As an overall tendency, there wasn't that much difference by groups. In terms of perks for which people thought it acceptable to disclose personal data, it has been proved that more people thought it acceptable to disclose only tangible goods such as “money/point” and “goods and services” as a whole. Without much difference in tendencies by groups, it has been proved that more people even in “faction of antipathy toward exploitation” and “faction of sensitivity” thought it acceptable to disclose personal data for tangible goods. In the future, it is important for companies to focus on how they collect as much information as consumers give consent to and how they apply the information to marketing, as well as clearly disclose how personal information is collected.

References

1. Mitani, K.: CRM marketing strategy. Toyo Keizai, Inc. (2003)
2. Nakagiri, D.: Web marketing and consumer behavio. Japan Society for Information and Management 29(3), 23–28 (2008)
3. Nishihira, K.: Marketing strategy in Internet space and methodology of consumer action research –Viewpoint of modern and postmodern. A Treatise on Electricity and Magnetism, Research of quotient study 27, 53–67 (2007)
4. Sakamoto, K.: Information strategy of Japan of the 21st century. Iwanami Shoten, Japan (2002)
5. Toyoda, H.: Kyobunsan kozo bunseki (kozo hoteishiki moderingu) [gijyutsu hen] [Covariance structure analysis (Structural equation modeling) [Technological chapter]] [summary]. Asakura Shoten Inc., Japan (2003)
6. Toyoda, H.: Kyobunsan kozo bunseki (kozo hoteishiki moderingu) [nyumon hen] [Covariance structure analysis (Structural equation modeling) [Introductory chapter]] [summary]. Asakura Shoten Inc., Japan (1998)

Modelling Social Cognitive Theory to Explain Software Piracy Intention

Ameetha Garbharran and Andrew Thatcher

School of Human and Community Development, University of the Witwatersrand, Johannesburg, South Africa
ameetha@e-pod.co.za, andrew.thatcher@wits.ac.za

Abstract. This study found evidence for a viable social cognitive model of software piracy intention using path analysis. Compared to outcome expectations, moral disengagement emerged as the stronger significant mediator in the model exerting its influence on the relationships between past behaviour and future intention and between past behaviour and outcome expectations. This study found weak evidence for the mediating influence of facilitators and impediments while their impact as a moderating variable was inconclusive.

Keywords: social cognitive theory, software piracy, path analysis, social cognitive model, software piracy intention.

1 Introduction

Software piracy is a phenomenon whose impact has pervaded the globe. Its appeal, however, has not been confined to the owners of the copyright of digital intellectual property or the groups of individuals that do not believe in the legitimacy of these rights. It has also captured the interest of psychologists who strive to grasp the essence of human behaviour. Recent endeavors yielded theories that sought to deconstruct software piracy intention and behaviour into their basic building blocks. Drawing on insights yielded by some of the more popular social psychological models, this paper proposes a social cognitive model for predicting software piracy intention. In this study software piracy is defined as the unauthorised copying or distribution of digital material [1] without permission from and with no intention to compensate the copyright holders for the use or dissemination of their intellectual property [2].

Social psychology's popular models of human behaviour such as the theory of reasoned action [3], [4], the theory of planned behaviour [5], [6], [7] the theory of interpersonal behaviour [8] and social cognitive theory [9], have examined the constructs that predict intentions to pirate software and software piracy behaviour. The core elements of attitudes, social pressures, intention and behaviour are consistent across the social psychological theories used to investigate software piracy. A significant improvement to the theory of reasoned action (TRA) was a formal acknowledgement of the environment as an influencer of human behaviour which was captured in the controllability dimension of perceived behavioural control in the theory of planned behaviour (TpB), facilitating conditions in the theory of interpersonal

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 301–310, 2011.

behaviour (TIB), and facilitators and impediments in social cognitive theory (SCT). While the TIB included affect and habit to understand software piracy intentions and behaviour, social cognitive research in this domain has not emphasised these constructs. As a theoretical paradigm, however, the personal factors construct as a primary determinant of human behaviour in SCT caters for the inclusion of variables such as habit and affective factors. A significant difference between SCT and the other social psychological theories is the incorporation of moral disengagement into its theoretical fabric.

The TpB was found to be superior to the TRA for predicting software piracy intention due to its inclusion of perceived behavioural control [5]. Social cognitive theory was predictively superior to the TpB in a study investigating intentions to download pirated music. The absence of the self-regulation process in the TpB (moral disengagement is situated in the self-regulation process in SCT), was one of the reasons it was considered inadequate for explaining the impact of norms on intentions to pirate music [9]. The inclusion of moral disengagement as an intrinsic part of SCT could elevate its predictive utility for explaining antisocial behaviour. Unlike the TpB onto which constructs such as moral obligation and past behaviour were tacked based on the context in which the theoretical framework was applied, moral disengagement, past behaviour, outcome expectations, self-efficacy, and facilitators and impediments are woven into the fabric of SCT as predictors of intention and behaviour. Of the social psychological theories of human behaviour, therefore, SCT is identified as one of the most comprehensive frameworks that explains a wide range of prosocial and antisocial behaviours without having to borrow concepts from other theories.

2 A Social Cognitive Model of Software Piracy

Social cognitive theory is a framework for understanding human behaviour in which personal factors, environmental influences and behaviour interact in a pattern of triadic reciprocal causation [10]. Bandura [10] was not explicit about the constituent components of SCT but research on its use in the promotion of healthy behaviours led to the explication of some of its core determinants including, self efficacy, outcome expectations, facilitators and impediments and intention (when behaviour is the outcome variable) [11], [12]. These constructs will be modelled alongside moral disengagement [10] to examine the factors that influence intention to pirate software in this study.

Past behaviour refers to the extent of enactment of a specific behaviour in the past. With behaviour posited as a key determinant of human thought and action, past behaviour is a firm predictor of future behaviour and is inextricably interwoven into SCT. Social cognitive theory defines self efficacy as “people’s judgments of their capabilities to organise and execute courses of action required to attain designated types of performances” [10:391]. Moral disengagement refers to the mechanisms individuals activate to override the influence of their internal self-sanctions and to distance themselves from perceived reprehensible consequences of their behaviour [10]. Social cognitive theory is the only framework that caters for the activation of moral disengagement in antisocial contexts. Outcome expectations refer to anticipatory judgments about the likely consequences of actions [10]. Bandura’s [10] definition of outcome expectations in SCT included attitudes and subjective norms as they have traditionally been defined in the theories of reasoned action and planned behaviour [13]. Facilitators and impediments in SCT encompass situational and contextual factors that render behaviour easier or more difficult to enact [10]. Intention refers to the likelihood of enacting a specific behaviour in the future. In SCT intention is defined as "the determination to perform certain activities or to bring about a certain future state of affairs" [10:467]. Intention is conceptualised as an immediate precursor to behaviour. In studies employing cross-sectional designs (such as this one) it is not feasible to declare behaviour as the dependent variable since all variables are measured concurrently which would imply that any information collected about behaviour would necessarily pertain to past behaviour. In these instances, intention is the logical dependent variable [14].

Empirically, previous enactments of software piracy have had meaningful direct effects on future intention to pirate software [7], [15]. Past behaviour could affect the nature and extent of subsequent moral disengagement. If individuals have pirated software before, it is likely that they will have previously invoked the moral disengagement mechanism. In subsequent enactments, while moral disengagement may still be necessary, it is possible that less moral disengagement is required to distance oneself from the reprehensible consequences of one's actions and/or that the format moral disengagement takes when one is engaging in an antisocial behaviour for the first time compared to when s/he engages in multiple instances of the same antisocial behaviour (or of antisocial behaviour, in general) could vary. Past behaviour could inform subsequent anticipatory judgments about the likely consequences of the same behaviour intended for enactment in the future. If previous enactments led to positive outcomes, it is likely that individuals will form positive outcome expectations about enacting the same behaviour in the future. This conceptualization allows past behaviour to precede moral disengagement and outcome expectations in the model.

There is uncertainty about whether self-efficacy precedes moral disengagement in SCT and some researchers have not accommodated moral disengagement and self-efficacy in the same model when explaining antisocial conduct [16]. In this study self-efficacy is regarded as an exogenous variable which is correlated with past behaviour (if one has successfully pirated software in the past one is likely to have an elevated belief in his/her proficiency to engage in the behaviour in the future). It would have been equally viable to position self-efficacy as an outcome of past behaviour with direct links to outcome expectations and intention and an indirect link to intention through outcome expectations and facilitators and impediments individually and in combination.

The empirical research showed support for direct effects between moral disengagement, outcome expectations and facilitators and impediments on future intention to pirate software [15]. This suggests that if individuals distance themselves from the reprehensible consequences of their actions; if they believe that pirating software will yield positive outcomes; and if they perceive conditions in the environment as enabling, they are more likely to form an intention to pirate software. Evidence for the effect of self-efficacy on intention, however, has been mixed. Garbharran and Thatcher [14] found that the effect of self-efficacy on intention may have been mediated by the facilitators and impediments construct while other research yielded a significant direct path between them [15], [17]. If individuals have the proficiency to pirate software it could influence their intentions to engage in the behaviour. However, simply believing in one's ability to enact a behaviour does not imply that one will necessarily execute it. Therefore, it is likely that expectations of positive outcomes envisaged from performing the behaviour could mediate the relationship between self-efficacy and intention. This relationship was investigated in a study by LaRose and Kim [9] who found that outcome expectations acted as a partial mediator of the relationship between self-efficacy and intention. It is also likely that the perception of enabling conditions in the environment could serve as an incentive to individuals who are adept at pirating software to form positive outcome expectations and, consequently, develop future piracy intentions.

The mediating role of moral disengagement has been explored in empirical research. However, there is inconclusive evidence for its effect as a mediator between the social cognitive predictors of self-efficacy, social norms and attitudes and intention [17]. A test for its role as a moderator of the relationships between self-efficacy, attitudes and social pressures on intention revealed no moderating effect [18]. This study hypothesises that moral disengagement and outcome expectations, individually and in combination, mediate the relationship between past behaviour and intention, that moral disengagement mediates the relationship between past behaviour and outcome expectations and that outcome expectations mediate the relationship between moral disengagement and intention. The argument for the mediating influence of moral disengagement on the relationship between past behaviour and outcome expectations is individuals who have engaged in software piracy in the past are first required to distance themselves from the egregious consequences of their prospective actions before being able to form positive anticipatory judgments about the consequences of engaging in the same behaviour in the future. The mediating influences of outcome expectations and moral disengagement individually on the relationship between past behaviour and intention suggest that if individuals have pirated software before, they need to form positive anticipatory judgments about the likely consequences of their choices first before developing future intentions to pirate software; and if individuals have pirated software previously the extent to which they are able to distance themselves from the negative consequences of their intended behaviour is likely to inform their future intentions. The mediating effect of moral disengagement and outcome expectations in combination on the relationship between past behaviour and intention suggests that if individuals have pirated software before, they first need to morally disengage in order to develop positive expectations of the outcomes of their intended behaviour before forming future intentions to pirate.

External environmental effects have directly affected actual piracy behaviour in previous research [8]. In the context of a cross-sectional research design where intention is the dependent variable, environmental influences are hypothesised to have a direct effect on piracy intention. In the model under construction in this study, facilitators and impediments are further believed to mediate the effects of moral disengagement, outcome expectations and self-efficacy on intention. If individuals morally disengage from the negative consequences of a behaviour, anticipate positive outcomes from engaging in the behaviour and/or perceive they have the efficacy to enact the behaviour, factors in the environment which either act as enablers or inhibitors to the formation of software piracy intentions could mediate these relationships.

This exploratory study is geared towards answering the following research questions:

1. Does the social cognitive model offer a viable explanation of software piracy intention?
2. Do moral disengagement and outcome expectations act as mediators in the model?
3. Do facilitators and impediments mediate or moderate paths in the model?

3 Method

Sample. The sample consisted of a group of 456 professionals from a range of industry sectors in South Africa. There was a larger proportion of males (77%) than females (23%) in the sample. Individuals between the ages of 35 and 54 made up 55% of the sample while individuals between the ages of 18 and 34 comprised 20% and the remaining 25% was made up of individuals aged 55 and above. The largest proportion of individuals in the sample (48%) achieved formal university qualifications (undergraduate: 15%; postgraduate: 33%) and worked in the information technology sector (43%).

Procedure. Data was collected using an online survey tool. Emails, with a link to the online survey, were sent to prospective respondents from a database of consumers of technology-oriented products and services from a large telecommunications organisation inviting them to complete the survey. Invitations were followed by a reminder email. Data collection spanned two and half months.

Measurements. A questionnaire was developed to measure the social cognitive constructs for explaining software piracy intention. It consisted of 27 items clustered into six discrete scales for past behaviour, moral disengagement, self-efficacy, outcome expectations, facilitators and impediments and intention. Items for self-efficacy, outcome expectations, moral disengagement, facilitators and impediments and intention were measured using a 5-point Likert-type scale which ranged from strongly disagree (1) to strongly agree (5). The 5-point Likert-type rating scale for past behaviour ranged from never (1) to very often (5).

Analysis. The data were analysed using the CALIS procedure in SAS9.2. A path analysis was conducted to test the fit of the model to the data and to derive estimates of the hypothesised paths. Direct and indirect effects were analysed to detect the presence of mediator variables and multiple-group path analysis was performed to test for the presence of a moderator.

4 Results

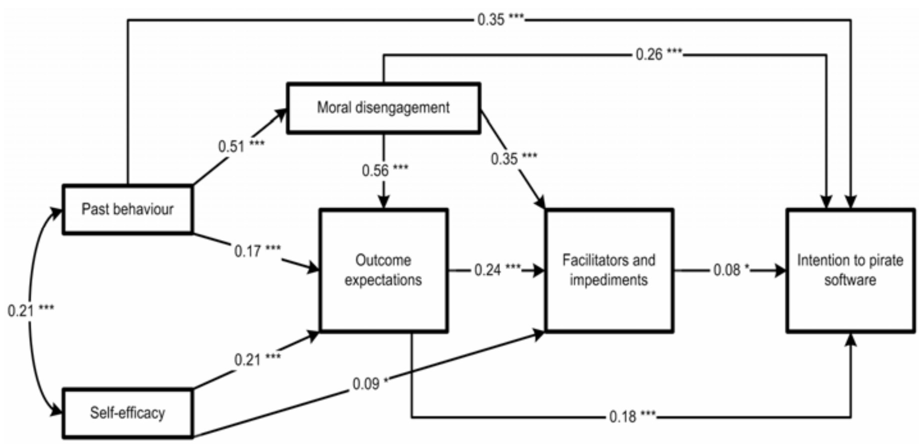
Internal reliabilities were measured using Cronbach's coefficient alpha and are reported in Table 1 together with means and standard deviations.

Table 1. Reliability estimates, means, standard deviations and intercorrelations

	Variable	Items	Cronbach's α	Mean	SD	1	2	3	4	5	6
1	Intention	3	0.86	6.08	2.94	1.00					
2	Facilitators & impediments	6	0.61	16.65	3.97	0.42**	1.00				
3	Outcome expectations	4	0.71	2.63	0.80	0.57**	0.50**	1.00			
4	Moral disengagement	8	0.87	20.23	7.05	0.60**	0.51**	0.66**	1.00		
5	Self-efficacy	3	0.86	3.38	1.02	0.19**	0.18**	0.28**	0.07	1.00	
6	Past behaviour	3	0.83	1.09	1.22	0.60**	0.32**	0.50**	0.51**	0.21**	1.00

** p<0.001

The proposed model was evaluated using the maximum likelihood estimation technique and the covariance matrix was analysed. The overall model fit indices suggested a good fit to the data ($\chi^2_{(2,\ N=456)}$=0.084, p=0.646; SRMSR=0.011; RMSEA=0.000 [90% CI=0.00–0.073], CFit=0.858; GFI=0.999; CFI=1.000; Critical N=3118). However, the residuals and modification indices hinted at points of strain. The Wald test suggested that the direct path from self efficacy to intention could be dropped without negatively impacting overall model fit. The chi-square difference test ($\chi^2_{\mathrm{diff}(1,\ N=456)}$=0.77, p=0.858) revealed no significant decrement in model fit from eliminating the direct path between self-efficacy and intention. Therefore, this path was dropped to yield the final model tested in this study; depicted in Figure 1.



* p<0.05 *** p<0.001

Model fit summary: $\chi^2_{(3,\ N=456)}$ = 1.640, p=0.650; SRMSR = 0.011; RMSEA = 0.000 [90% CI = 0.00 – 0.062], CFit = 0.898; GFI = 0.999; CFI = 1.000; Critical N = 2168.

Fig. 1. Path coefficients for a social cognitive model of software piracy intention

An evaluation of overall goodness of fit ($\chi^2_{(3,\ N=456)}$=1.640, p=0.650; SRMSR=0.011; RMSEA=0.000 [90% CI=0.00–0.062], CFit=0.898; GFI=0.999; CFI=1.000; Critical

N=2168) of the final model suggested that it fitted the data well. The residuals and modification indices did not indicate localised areas of strain. The path coefficients (depicted in Figure 1) were statistically significant and in the predicted direction [19]. The model explained 51% of the variance in intention (the main dependent variable).

To test the hypothesis of whether facilitators and impediments acted as a moderator, a multiple-group path analysis was conducted. The sample was split into two groups based on perceptions of high and low facilitating conditions in the environment, to assess changes in the relationships between past behaviour and intention, moral disengagement and intention and outcome expectations and intention at different levels of facilitators and impediments. While the findings presented in Table 2 suggested that the model with facilitators and impediments as a moderator did not offer an adequate fit to the data, preliminary analyses revealed that facilitators and impediments seemed to be a significant moderator of the relationships between past behaviour and intention and moral disengagement and intention. However, it did not appear to moderate the relationship between outcome expectations and intention as hypothesised. Instead a significant moderating effect for facilitators and impediments was found between past behaviour and moral disengagement.

Table 2. Tests for the effect of facilitators and impediments as a moderating variable

	χ^2	*df*	χ^2_{diff}	df_{diff}	*p*
Null model: All parameters constrained	43.99	17			
Past behaviour ⇨ Intention released	38.17	16	5.82	1	0.02
Past behaviour ⇨ Moral disengagement released	39.56	16	4.44	1	0.04
Moral disengagement ⇨ Intention released	39.70	16	4.29	1	0.04
Outcome expectations ⇨ Intention released	42.43	16	1.56	1	0.21
Past behaviour ⇨ Intention + Past behaviour ⇨ Moral disengagement + Moral disengagement ⇨ Intention released	32.54	14	11.45	3	<0.01

Note: The sample was split into two groups based on high and low scores on the facilitators and impediments construct to test its effect as a moderating variable.
Overall model fit summary: $\chi^2_{(14, N=456)} = 32.54$, $p<0.003$; SRMSR = 0.093; RMSEA = 0.076 [90% CI = 0.421 – 0.111], CFit = 0.095; GFI = 0.975; CFI = 0.972; Critical N = 332
Model 1 (Low F+I) fit summary: SRMSR = 0.085; GFI = 0.987; NFI = 0.976; 26% contribution to χ^2
Model 2 (High F+I) fit summary: SRMSR = 0.102; GFI = 0.961; NFI = 0.928; 74% contribution to χ^2

5 Discussion

The social cognitive model investigated in this study is viable and offers one possible explanation of intention to pirate software. The model investigated in this study was one of two believed to offer equally plausible explanations of intention to pirate software. In the equivalent model, self-efficacy was an endogenous variable with a direct relationship to past behaviour. Both models fit the data equally well and the authors leave it to the reader to decide which theoretical interpretation is more meaningful.

A significant indirect effect was noted between past behaviour and outcome expectations (direct=0.18, p<0.001; indirect=0.29, p<0.001). In this study it was hypothesised that moral disengagement mediated this relationship and the findings confirmed that this construct was a significant partial mediator of the relationship between past behaviour and outcome expectations. This implies that the enactment of software piracy in the past does not presuppose the automatic expectation of positive outcomes for engaging in the same behaviour in the future. Prior software piracy behaviour presupposes that the moral disengagement mechanism was activated previously. This finding suggests that despite its prior activation, further moral disengagement may be necessary in order for individuals who have pirated software in the past to form positive outcome expectations about engaging in the behaviour again in the future.

The relationship between past behaviour and future intention was characterised by a significant indirect effect (direct=0.35, $p<0.001$; indirect=0.24, $p<0.001$) suggesting that moral disengagement, outcome expectations and facilitators and impediments acted as partial mediators in the model. These findings imply that in order for individuals who have pirated software in the past to form intentions to pirate software again in the future, it is necessary for them to activate the moral disengagement mechanism, to consider software piracy an act that will lead to beneficial outcomes and to believe that external environmental forces will aid their intended behaviour. The indirect effect of moral disengagement (0.13, $p<0.001$) and outcome expectations (0.10, $p<0.01$) on the relationships between past behaviour and intention revealed that moral disengagement was more statistically significant. This suggests that moral disengagement is likely to be a more meaningful mediator of the relationship between past behaviour and future intention.

The effect of self-efficacy on future intention appeared to be completely mediated by the facilitators and impediments construct (direct=0.03, $p<0.38$; indirect=0.05, $p<0.001$) while the indirect effect of outcome expectations on the relationship between self-efficacy and intention was not significant (direct=0.17, $p<0.001$; indirect=0.02, $p<0.07$). The direct path between self-efficacy and intention was not significant and was consequently dropped from the model. However, its effect through the facilitators and impediments variable was significant but weak. This study did not find a strong or significant direct path between self-efficacy and intention to pirate software suggesting that one's belief in one's efficacy to engage in the behaviour did not automatically translate into software piracy intentions. Instead, the effect of self-efficacy on intention was mediated through facilitators and impediments to influence intentions suggesting that if one had the efficacy to engage in a behaviour, one was more likely to form intentions to enact that behaviour if the environment presented enabling conditions.

Facilitators and impediments appeared to mediate the relationship between self-efficacy and software piracy intention. Its role as a mediator was significant but weak. Traditionally, environmental influences moderate the relationships between variables of theoretical interest. This implies that relationships are likely to be impacted by different levels of facilitators and impediments in the environment. It seems that individuals who pirated software in the past are more likely to develop future intentions to pirate software if there are strong environmental enablers to support their intended actions. It also seems that morally disengaging from negative consequences of the behaviour is more likely to lead to the formation of software piracy intentions in the context of strong environmental enablers. Based on the findings in this study, individuals who pirated software in the past were more likely to morally disengage from the reprehensible consequences of their actions when strong environmental facilitators supported the enactment of the intended behaviour. However, it seems counterintuitive that positive outcome expectations led to intentions to pirate software regardless of the level of environmental enablers as it would be expected that the expectation of positive outcomes from pirating software would lead to the formation of software piracy intentions in the context of strong environmental enablers. This unexpected finding could be an artifact of a construct with poor reliability and/or a poor fitting model and requires corroboration.

6 Limitations and Directions for Future Research

This study was cross-sectional and could only offer tentative explanations of causality among the social cognitive variables. A longitudinal research design in future research endeavors could better clarify questions of causality and temporal precedence. In addition, a longitudinal design would cater for the inclusion of behaviour as the primary outcome variable. It is possible that facilitators and impediments could moderate the relationship between intention and behaviour such that higher levels of facilitators in the environment will translate into piracy behaviour if intentions to pirate software were formed whereas low levels of enablers in the environment could inhibit software piracy behaviour even if individuals form intentions to engage in the behaviour. This should be explored in future research.

7 Conclusion

The social cognitive model offered a viable explanation of intention to pirate software rendering it a useful theoretical framework to explain software piracy; one instance of antisocial behaviour. Moral disengagement, SCT's unique contribution to explaining antisocial behaviour, was identified as the strongest mediator between past behaviour and outcome expectations in the model highlighting the necessity for further moral disengagement in individuals who had previously engaged in software piracy in order to form positive outcome expectations about the possibility of engaging in the behaviour in the future. This suggested that moral disengagement may not be a once-off process even when the envisaged future behaviour is identical in principle to past behaviour. Thus, while the nature and/or extent of moral disengagement may alter over time, it does not become redundant in the context of multiple enactments of the same behaviour. Moral disengagement was a more significant mediator than outcome expectations of the relationship between past behaviour and future piracy intention. Weak yet significant evidence for the mediating effect of facilitators and impediments in the proposed model was found. Even though this study was unable to comment conclusively on the moderating role of facilitators and impediments in a social cognitive model of software piracy intention, preliminary findings suggested that facilitators and impediments may be more optimal as moderators of the relationships between social cognitive constructs.

Acknowledgments. We thank Michael Greyling from the Department of Psychology at the University of the Witwatersrand for his expert statistical input. The financial assistance of the National Research Foundation (NRF) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at are those of the authors and may not necessarily be attributed to the NRF.

References

1. Business Software Alliance, `http://www.bsa.org/country.aspx?sc_lang=en-ZA`
2. Gopal, R.D., Sanders, G.L.: International software piracy: Analysis of key issues and impacts. Info. Sys. Research 9(4), 380–397 (1998)
3. Eining, M.M., Christensen, A.L.: A psycho-social model of software piracy: the development and test of a model. In: Dejoie, R., Fowler, G., Paradice, D. (eds.) Ethical Issues in information systems, pp. 182–188. Boyd & Fraser, Boston (1991)
4. Woolley, D.J., Eining, M.M.: Software piracy among accounting students: A longitudinal comparison of changes and sensitivity. J. Info. Sys. 20(1), 49–63 (2006)
5. Chang, M.K.: Predicting unethical behavior: A comparison of the theory of reasoned action and the theory of planned behavior. J. Bus. Eth. 17, 1825–1834 (1998)
6. Peace, A.G., Galletta, D.F., Thong, J.Y.L.: Software piracy in the workplace: A model and empirical test. J. Mngt. Info. Sys. 20(1), 153–177 (2003)
7. Cronan, T.P., Al-Rafee, S.: Factors that influence the intention to pirate software and media. J. Bus. Eth. 78, 527–545 (2008)
8. Limayem, M., Khalifa, M., Chin, W.W.: Factors motivating software piracy: a longitudinal study. IEEE Trans. Eng. Mngt. 51, 414–425 (2004)
9. LaRose, R., Kim, J.: Share, steal or buy? A social cognitive perspective of music downloading. CyberPsych. & Beh. 10(2), 267–277 (2007)
10. Bandura, A.: Social foundations of thought and action: A social cognitive theory. Prentice-Hall, Englewood Cliffs (1986)
11. Bandura, A.: Self-efficacy: The exercise of control. W.H. Freeman and Company, New York (1997)
12. Bandura, A.: Health promotion by social cognitive means. Health Educ. & Beh. 31(2), 143–164 (2004)
13. Ajzen, I.: The theory of planned behavior. Org. Beh. & Hum. Proc. 50, 179–211 (1991)
14. Garbharran, A., Thatcher, A.: A case for using a social cognitive model to explain intention to pirate software. J. eHealth Tech. & App. 7(2), 87–98 (2009)
15. Garbharran, A., Thatcher, A. (under review): Predicting software piracy intentions in organisations: Does organisation size matter? J. Bus. Eth. (2010)
16. Bandura, A., Barbaranelli, C., Caprara, G.V., Pastorelli, C.: Mechanisms of moral disengagement in the exercise of moral agency. J. Pers. & Soc. Psych. 71, 364–374 (1996)
17. Matthews, M.S.: Bandura's social cognitive theory in a cross-nation study of software piracy. Unpublished masters research report. University of the Witwatersrand, Johannesburg, South Africa (2009)
18. Wentzell, A.: The implementation of social cognitive theory in the understanding of the unauthorised copying of software. Unpublished masters research report. University of the Witwatersrand, Johannesburg, South Africa (2008)
19. Brown, T.A.: Confirmatory Factor Analysis for Applied Research. The Guilford Press, New York (2006)

A Practical Analysis of Smartphone Security*

Woongryul Jeon[1], Jeeyeon Kim[1], Youngsook Lee[2], and Dongho Won[1,**]

[1] School of Information and Communication Engineering, Sungkyunkwan University, Korea
wrjeon@security.re.kr, jeeyeonkim@paran.com,
dhwon@security.re.kr
[2] Department of Cyber Investigation Police, Howon University, Korea
ysooklee@howon.ac.kr

Abstract. Recent developments in mobile technologies have produced a new kind of device, a programmable mobile phone, the smartphone. Generally, smartphone users can program any application which is customized for needs. Furthermore, they can share these applications in online market. Therefore, smartphone and its application are now most popular keywords in mobile technology. However, to provide these customized services, smartphone needs more private information and this can cause security vulnerabilities. Therefore, in this work, we analyze security of smartphone based on its environments and describe countermeasures.

Keywords: Smartphone, Smartphone security, Security analysis, Security mechanism.

1 Introduction

A smartphone is a mobile phone that offers more advanced computing ability and connectivity than a contemporary basic feature phone. Smartphones and feature phones may be thought of as handheld computers integrated within a mobile telephone, but while most feature phones are able to run applications based on platforms such as Java ME, a smartphone allows the user to install and run more advanced applications based on a specific platform. Smartphones run complete operating system software providing a platform for application developers[1].

Based on this feature, smartphone user can develop any programs which are customized in specific needs, and this is a most powerful advantage of smartphone. For example, smartphone user can search most popular restaurant, or nearest bus stop. Furthermore, smartphone user can trade their assets like stocks or use banking service with wireless network. Smartphone user can send or receive e-mails, too.

However, to provide these services, smartphone needs more private information than feature phone, thus, it is very important to keep smartphone secure. If smartphone user

* This research was supported by Basic Science Research Program thorough the National Research Foundation of Korea(NRF) funded by the Ministry of Education, Science and Technology(2010-0011279).

** Corresponding author.

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 311–320, 2011.

lost his/her smartphone, for example, every information like address, e-mail, log data in web browser, SMS(Short Message Service), MMS(Multi Message Service) or others, can be exposed if there is no appropriate security solutions.

In present, there are many researches on smartphone security, but there is lack of effort to analyze all security threats of smartphone. To establish smartphone security, security threats based on smartphone environment is necessary. This work is needed for design of security solution, either, to prevent potential vulnerabilities of smartphone.

Therefore, in this paper, we analyze security of smartphone and suggest countermeasures. This paper consists of five sections. In section two, we analyze smartphone environment and its assets. In section three, we analyze security of smartphone and derive vulnerabilities and threats of smartphone. In Section 4, we describe applicable security mechanisms for smartphone, and finally, in section 5, we end with a conclusion.

2 Security Environments of Smartphone

2.1 Environments of Smartphone

Smartphone can be connected various subjects, internet, PC, other mobile devices using wireless network. This feature makes smartphone useful and most popular mobile device. However, in other words, this feature means that malicious attacker or software can invade smartphone in various paths. Following figure 1 shows general environment of smartphone.

Fig. 1. Environments of Smartphone

The user can make a call or receive a call, manage his/her schedules, play game or use other functions by his/her smartphone. Some applications may need to connect web or other devices to provide customized service, and in this case, smartphone can connect others with various wireless network technologies.

The base station is kind of way to connect web and it is a basis of calling service. The base station relays phone calls, messages, e-mails or various data via 3G networks. If there is AP(Access Point)s around user smartphone, user also can connect to web using AP.

The satellite provide location information of smartphone, and this information can be used various services, for example, map, messenger, even if when user take a picture, location information is inserted in picture.

The PC(Personal Computer) can be connected to smartphone by cable or wireless network, and user may download files or update firmware through PC.

These entities can be regarded as both target and mean of attack. For example, an attacker can take denial of service attack on base station or web server and risk availability of smartphone. Furthermore, some entities like web server or PC can be used as a host to infect user smartphone with malware. Therefore we have to consider these entities in smartphone security.

2.2 Assets of Smartphone

Now we have to define smartphone assets, because the asset can be regarded by target of attack, and threats and vulnerabilities are basis of the attack. Following table 1 shows assets of smartphone in this paper.

Table 1. The Assets of Smartphone

Assets	*Description*
Private Information	Address book, Calling history, Location information, Notebook, Schedule, Cache file of web browser, password used in web, email and its attachments, and other information
Device	Smartphone device System resources of smartphone(CPU, RAM, Battery or etc.)
Applications	Smartphone applications user installed

First, information in smartphone can be defined an asset of smartphone. The information include all data both stored in smartphone and transmitted out to smartphone, for example, address book, calling history, location information, e-mail and its attachments, SMS(short message service), media files and so on forth. These information is managed by applications of smartphone, thus for security of smartphone, the application is an essential entity.

Second, smartphone itself can be defined an asset. Because smartphone can make a call or connect wireless network, thus malicious user who get smartphone someone lost, can cause overcharging by using smartphone. In addition, resources of smartphone can be regarded as an asset, because these resources ensure availability of smartphone. In fact, some malwares exhaust resource of smartphone on purpose to risk availability.

Third, applications on smartphone can be defined an asset. There are two kinds of applications, one is freely distributed by user or online application store, and another is commercially used with digital rights. The smartphone user has to pay some charge to use commercial applications and thus, application itself can be regarded as an asset.

Furthermore, the applications are closely related the information, thus it is natural to regard the application to an asset of smartphone. For example, most of smartphone web browser stores user's ID and password which can be used in online authentication process. Generally, smartphone provides QWERTY pad to input device and this device is implemented in touch screen about 3~4 inches, so, it is inconvenience to type ID and password every times. Therefore web browsers on smartphone store ID and password, and this feature is the reason why applications have to be regarded as smartphone assets.

3 Vulnerabilities and Threats of Smartphone

In this section, we derive vulnerabilities and threats of smartphone. All vulnerabilities respond specific threat.

3.1 Vulnerabilities of Smartphone

When to keep a system secure, we have to consider how keep system secure, and answer of this question is what kinds of threats can harm the system. To identify all existing threats, first, we have to confirm the assets and this is described above section. In this section, we determine security objectives for the system. In principle, the security of mobile devices deals with the same issues conventional computer security deals with confidentiality, integrity and availability. Table 2 shows the security objectives in this paper.

Table 2. Security Objectives

Issues	*Description*
Confidentiality	Confidentiality determines who is allowed to access what.
Integrity	Integrity identifies who is allowed to modify or use a certain resources.
Availability	Availability describes the requirement that a resource be usable by its legitimate owner.

Table 3. Vulnerabilities of Smartphone

<table>
<tr><th colspan="2">Vulnerabilities</th><th>Description</th></tr>
<tr><td rowspan="3">Internal of Smartphone</td><td>V1. Implementation error</td><td>Malfunction caused by implementation error
Malicious attacker can take advantages using implementation error(e.g. type safety, arbitrary code execution)</td></tr>
<tr><td>V2. Incompatibility</td><td>Disabling application caused by incompatibility between applications
Disabling application caused by incompatibility between application and platform</td></tr>
<tr><td>V3. User unawareness</td><td>Unawareness of a device owner to risks of installing applications from un-trusted sources
Unawareness of the device owner to the risk of connecting to un-trusted Wi-Fi networks and web-sites(e.g. Rogue AP, Phishing Site).
Unawareness of the device owner to the risks posed by improper configuration (e.g., Bluetooth settings, browser settings).
Unawareness of the owner to social engineering attacks.
User can lose his/her smartphone.</td></tr>
<tr><td rowspan="2">External of Smartphone</td><td>V4. Vulnerabilities of Wireless Network</td><td>Corrupting, blocking or modifying information on the wireless network by sniffing, spoofing or eavesdropping</td></tr>
<tr><td>V5. Vulnerabilities of External Objects</td><td>External objects of smartphone environment like web server, AP, base station or PC can risk smartphone security by its potential vulnerabilities, insecure management, or so on forth.</td></tr>
</table>

The threats can be divided in two groups, vulnerability and threat. Vulnerability means that it can risk security object potentially, and threat means that it can risk security object directly. In this paper, threats and vulnerabilities can be described based on this form, for example, resident malware can alter smartphone configuration without authority. This example includes subject – malware, object – system configuration, and action – altering. Now, we can derive vulnerabilities and threats of smartphone.

Table 3 shows vulnerabilities of smartphone. V stands Vulnerability in the table [2-11].

3.2 Threats of Smartphone

Threats of smartphone give shape to attack using vulnerabilities. In this paper, we divide threats in two groups, Threats caused by attackers and Threats caused by user unawareness or intention.

Table 4 shows threats of smartphone[2-11].

Table 4. Threats of Smartphone

Threats	*Description*	*Vulnerability*
	Threats caused by attackers	
T1. Malware	Malware can alter or expose private information in smartphone Malware can risk availability by meaningless operation(e.g. arbitrary code execution) Malware can abuse costly services and functions(e.g. sending SMS/MMS, connecting wireless network)	V1 V3 V5
T2. Wireless Network Attack	An attacker can corrupt, block or modify information on the wireless network by sniffing, spoofing or eavesdropping	V4
T3. Denial of Service	An attacker can risk availability of smartphone to take denial of service attack to base station, wireless network, web server An attacker can risk availability of smartphone using radio interference	V4 V5
T4. Break-in	An attacker can gain partial or full control over the target smartphone by using flaw of code, code injection or abuse of logic error	V1
	Threats caused by user unawareness or intention	
T5. Malfunction	The user can disable or malfunction his/her application by mistake or misappropriate configuration Smartphone application can malfunction by incompatibility between platform and application.	V2 V3
T6. Phishing	The user can expose his/her private information by accessing phishing site The user can expose his/her private information by messenger phishing The user can expose his/her private information by SMS phishing	V3
T7. Loss	The user can lose his/her smartphone	V3
T8. Platform Alteration	The user can alter his/her smartphone platform intentionally(e.g. jail breaking in iPhone, rooting in android phone)	V3

Figure 2 presents the results of a qualitative risk analysis that we conducted in order to identify and prioritize the threats. This figure is based on the report published in December, 2010, by ENISA (European Network and Information Security Agency)[11]. According to this report, main risk of smartphone is user unawareness. Although most smartphone applications have privacy settings for controlling how and when location data is transmitted, but many users are unaware that the data is being transmitted or data can be hided, even many users are unaware of existence of the privacy setting to prevent this. Other likely threat is malware. Smartphone user can install malware to his/her smartphone by unawareness, SPAM mail, SMS, MMS or other various ways. Lost or stolen smartphone is one of the main threats. However, the report cluster this threat in unlikely and medium, because only 2% of smartphone user are lost or stolen their smartphone last year. However, smartphone is a small and light mobile device, so user usually loses his/her smartphone, and when it lost, whole of information can be exposed.

Intentional platform modification also can cause security problem. However, according to this report, only 10% of iPhone users unlock their smartphone called by jail breaking.

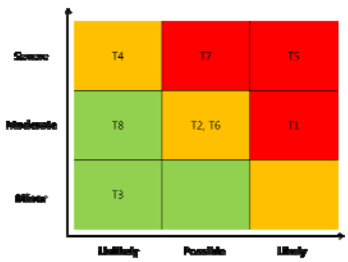


Fig. 2. Qualitative Threats Analysis

4 Applicable Security Mechanisms for Smartphone

Several security companies have already announced some security solutions for smartphone. These solutions include various antivirus software and intrusion detection systems that run on the smartphone and smartphone user can take these applications in online market. These applications can prevent attacks from outside like malware, but they can't prevent attacks from inside caused by using implementation error or user unawareness. Therefore to keep smartphone secure, it is required to adopt other security mechanisms, for example, platform modification, regular update and so on forth.

When considering the applicability of a security measure, we have to determine who would be implementing it and how should be realized. According to this, we can cluster smartphone security measures into three types in terms of their realization approaches[12]. Following table 5 shows definition of security mechanisms.

Table 5. Types of Approach

Types	*Description*
System Modification	Require altering platform's core source-code including the kernel **Advantage :** available new functionalities **Disadvantage :** Relatively expensive in terms of man power and time
System add-on	Require modification of platform's core configuration file **Advantage :** more easier than system modification **Disadvantage :** To adopt this modification, smartphone user have to re-install all applications
Add-on Applications	Can be applied by any user by simply installing an application **Advantage :** easy to adopt **Disadvantage :** If user does not install this application, there is no improvement in security

Add-on application is easiest way, however, in this way, to improve smartphone security, smartphone users have to install appropriate applications to their smartphone. Thus, this way can't ensure security improvement.

System add-on means system updates, and platform manufacturer can improve functionality and security in this way. This way also needs user-self update, but updates are perfectly adopted in new smartphone.

System modification is most expensive way to improve smartphone security, because it needs kernel configuration. However, this way can improve entire security of smartphone platform.

Table 6 shows applicable security mechanisms for smartphone.

T8, Platform alteration, potentially, can cause various security problem, thus in the table 6, we denote it in parentheses.

Adopting security solutions like anti-virus or SPAM filtering from appstore is easiest way to improve smartphone security, however, to adopt this way, smartphone user have to install applications. There are many applications for smartphone security, thus to improve smartphone security, the user should install necessary security solution.

In addition, smartphone users have to update their smartphone and applications periodically. Platform manufacturer and application developer provide updates for their products and this update includes both improvement of functionality and security. So, the user may update their smartphone platform and applications for smartphone security.

To ensure confidentiality and integrity in smartphone, application developer and smartphone user can adopt cryptographic technology. Cryptographic technology can be implemented two types, application and APIs. In application store, there are many applications using cryptographic technology. Some application provides data encryption for data confidentiality and some application provides hash function for data integrity, thus smartphone user can keep their smartphone secure using these applications. APIs also are provided to application developer, for example, several smartphone OS includes security library thus the developer can use these APIs in implementation.

Table 6. Applicable Security Mechanisms

Mechanisms	*Types*	*Description*	*Related Threats*
Anti-Virus Solution	Add-on Application, System Add-on	Anti-virus solutions scan files, memory, SMS, MMS, emails and URLs Anti-virus solutions can prevent malwares and also prevent access to phishing site	T1, T6, (T8)
Firewall	System Modification	Firewall blocks and/or audit un-allowed connections from/to device Firewall can prevent network attacks by denying access to untrusted wireless network	T3, (T8)
Secure API	System Add-on	Secure API provide cryptographic functionalities for application developer Application developer can implement secure functionality using secure APIs	T1, T2, (T8)
Access Control	System Modification	Access control limits access of processes and user to resources and/or services Access control can limit risk from malicious/exploited application	T1, T7
Authentication	System Modification	User should be authenticated to use device Authentication process can prevent unauthorized use of device	T7
Spam Filter	System Add-on, Application Add-on	SPAM filtering applications blocks MMS, SMS, emails and calls from unwanted origin SPAM filtering applications can prevent SPAM	T1
Pre-Testing	System Modification	Pre-Testing guarantee applications and authorizes developer Pre-Testing can prevent malware and ensure security of applications	T1, T4, T5
Regular Update	System Modification	Regular update for platform and smartphone application	T5
Remote Access Control	System Modification	Remote access control includes remote configuration and management of smartphone(remote blocking, remote reset) When user lose his/her smartphone, remote access control can reduce damage by lost smartphone	T7

In present, access control model has been studied in many papers, and this technology can provide advanced user authentication[9]. According to this model, multiple users can be classified in groups by his/her rights and access rights can be determined attributes of each group. This model is based on platform of smartphone. However, many smartphone users take root permission by jail breaking or rooting, thus access control model should consider this situation.

Remote access control is also studied and adopted in many ways. Because smartphone is a small and tiny device, smartphone users usually can lose their smartphone. When users lose their smartphone, remote access control is necessary to prevent exposure of data in smartphone and illegal usage of smartphone. Remote access control includes remote locking smartphone and remote reset of smartphone.

5 Conclusion

A smartphone is a mobile phone that offers more advanced computing ability and connectivity than a contemporary basic feature phone. Smartphones and feature phones may be thought of as handheld computers integrated within a mobile telephone, but while most feature phones are able to run applications based on platforms such as Java ME, a smartphone allows the user to install and run more advanced applications based on a specific platform. Smartphones run complete operating system software providing a platform for application developers.

Based on this feature, smartphone user can develop any programs which are customized in specific needs, and this is a most powerful advantage of smartphone. For example, smartphone user can search most popular restaurant, or nearest bus stop. Furthermore, smartphone user can trade their assets like stocks or use banking service with wireless network. Smartphone user can send or receive e-mails, too.

However, to provide these services, smartphone needs more private information than feature phone, thus, it is very important to keep smartphone secure. If smartphone user lost his/her smartphone, for example, every information like address, e-mail, log data in web browser, SMS(Short Message Service), MMS(Multi Message Service) or others, can be exposed if there is no appropriate security solutions.

In present, there are many researches on smartphone security, but there is lack of effort to analyze all security threats of smartphone. To establish smartphone security, security threats based on smartphone environment is necessary. Therefore, in this work, we analyzed security of smartphone and described applicable security mechanisms against threats.

References

1. Mulliner, C.R.: Security of Smart Phone, Master's Thesis of University of California (June 2006)
2. Guo, C., Wang, H.J., Zhu, W.: Smart-Phone Attacks and Defenses. In: HotNets III (November 2004)
3. Chen, J.V., Yen, D.C., Chen, K.: The acceptance and diffusion of the innovative smart phone use: A case study of a delivery service company in logistics. Information & Management 46(4) (2009)

4. Shabtai, A., Fledel, Y., Kanonov, U., Elovici, Y., Dolev, S.: Google Android: A State of the Art Review of Security Mechanisms, arXiv 2009 (November 2009)
5. http://seriot.ch/resources/talks_papers/iPhonePrivacy.pdf
6. http://www.mt.co.kr/view/mtview.php
7. Monk, A., Fellas, E., Ley, E.: Hearing only one side of normal and mobile phone conversations. Behaviour & Informaion Technology 23(5) (September 2004)
8. http://threatcenter.smobilesystems.com
9. Ni, X., Yang, Z., Bai, X., Champion, A.C., Xuan, D.: DiffUser: Differentiated User Access Control on Smartphones. In: Proc. 5th IEEE Int'l. Workshop on Wireless and Sensor Networks Security, WSNS 2009 (September 2009)
10. Schmidt, A.-D., Schmidt, H.-G., Clausen, J., Camtepe, A., Albayrak, S.: Enhancing Security of Linux-Based Android Devices. In: 15th International Linux Kongress (October 2008)
11. Shabtai, A., Fledel, Y., Kanonov, U., Elovici, Y., Dolev, S.: Google Android: A State-of-the-Art Review of Security Mechanisms, Cornell University library (2009)
12. Smartphone: Information security risks, opportunities and recommendations for users, ENISA Report (December 2010)

Cryptanalysis to a Remote User Authentication Scheme Using Smart Cards for Multi-server Environment*

Youngsook Lee[1,**], Jeeyeon Kim[2], and Dongho Won[2,***]

[1] Department of Cyber Investigation Police, Howon University, Korea
ysooklee@howon.ac.kr
[2] Department of Computer Engineering, Sungkyunkwan University, Korea
jeeyeonkim@paran.com,dhwon@security.re.kr

Abstract. Recently, Hsiang et al. proposed a remote user authentication scheme suited for multi-server environment, in which users can be authenticated anonymously using a smart card. This work reviews Hsiang et al.'s scheme and provides a security analysis on the scheme. Our analysis shows that Hsiang et al.'s scheme does not achieve its fundamental goal of not only any kind of authentication, either server-to-user authentication or user-to-server authentication but also password security. The contribution of the current work is to demonstrate these by mounting two attacks, a server impersonation attack and a user impersonation attack, on Hsiang et al.'s scheme. In addition, we demonstrate that their scheme is vulnerable to two-factor security which guarantees the security of the scheme when either the user's smart card or its password is stolen, but not both by employing the off-line dictionary attack.

Keywords: distributed system; authentication scheme; smart card; user anonymity; impersonation attack; off-line dictionary attack.

1 Introduction

In 2009, Hsiang et al. [2] proposed a remote user authentication scheme for multi-server environment in which users may send a login request message without disclosing their identities[5,7,11,3,9,8,10,13,1,6]. In addition to providing user anonymity, this scheme exhibits various other merits: (1) it allows the user to register only once with the registration center and then he/she is able to gain access to all servers included in multi-server environment without registering with every single server; (2) it does not require any server to maintain a password table for verifying the legitimacy of login users; (3) it allows users to choose and change their passwords according to their liking and hence gives more user

* This work was supported by Howon University in 2011.
** The first author.
*** Corresponding author.

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 321–329, 2011.

convenience; (4) it does not require synchronized clocks between in the network by using random numbers called nonces; (5) it is extremely efficient in terms of the computational cost since the protocol participants perform only a few hash function operations.

However, in this article, we uncover that Hsiang et al.'s scheme does not guarantee its main security goal authenticating between a remote individual and the server. We show these by mounting two attacks, a server impersonation attack and a user impersonation attack on Hsiang et al.'s scheme. Besides reporting the two security flaws, we demonstrate that the scheme is vulnerable to two-factor security [12] which guarantees the security of the scheme when either the user's smart card or its password is stolen, but not both by employing the off-line dictionary attack. What we do in this work is to report these security vulnerabilities of Hsiang et al.'s scheme.

The remainder of this paper is organized as follows. We begin by reviewing goal authenticating et al.'s remote user authentication scheme in Section 2. Then in Section 3, we present security weaknesses in goal authenticating et al.'s authentication scheme. Finally, we conclude this work in Section 4.

2 Review of Hsiang et al.'s Authentication Scheme

This section reviews a remote user authentication scheme proposed by Hsiang et al. [2]. The scheme participants include a registration center, a remote user, and multiple service provider servers. For simplicity, we denote the registration center by RC, the remote user by U_i, and the servers by S_1, S_2, ... , S_n. The scheme assumes that the registration center RC is a trust party responsible for securely delivering the secret keys to be shared with between U_i and S_j. Hsiang et al.'s scheme consists of two phases: registration phase and authentication phase. The registration phase is performed only once per user when a new user registers itself with the registration center. The authentication phase is carried out whenever a user wants to gain access to each server included in multi-server environment. Before the registration phase is performed for the first time, the registration center RC decides on the following system parameters: a one-way hash function h and three cryptographic keys x, y and r. The key y is shared securely among the registration center and all involved servers, while x and r are kept secret by the registration center.

2.1 Registration Phase

This is the phase where a new registration of a user takes place. The registration proceeds as illustrated by Fig. 1, where dashed lines indicate a secure channel:

Step 1. A user U_i, who wants to register with the registration center RC, chooses its password PW_i and a random number b at will. Then U_i computes $\pi_i = h(b||PW_i)$ and submits a registration request, consisting of its identity ID_i and π_i, to the registration center RC via a secure channel.

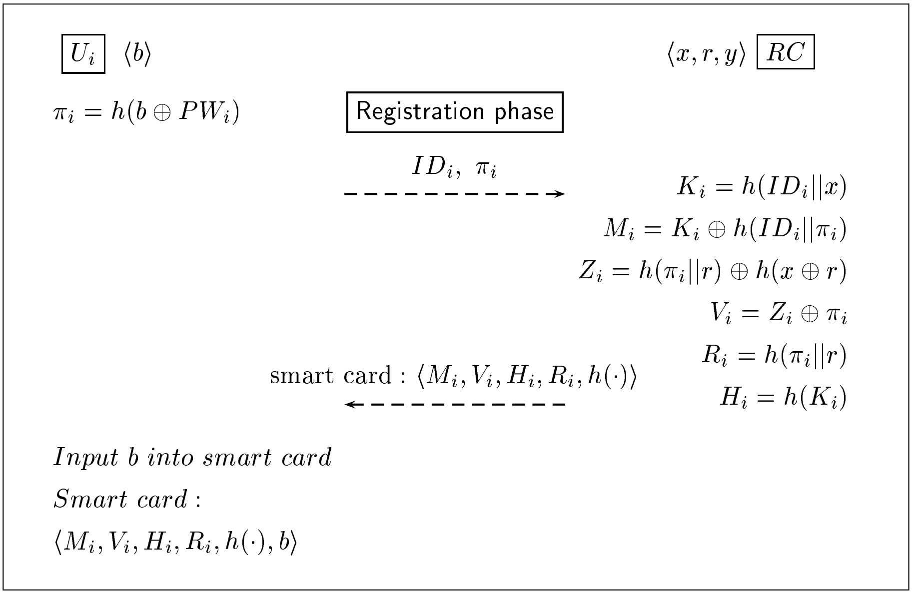


Fig. 1. The registration phase of Hsiang et al.'s scheme

Step 2. Upon receiving the request $\langle ID_i, \pi_i \rangle$, RC computes

$$\begin{aligned} K_i &= h(ID_i||x), \\ M_i &= K_i \oplus h(ID_i||\pi_i), \\ Z_i &= h(\pi_i||r) \oplus h(x \oplus r), \\ V_i &= Z_i \oplus \pi_i, \\ R_i &= h(\pi_i||r), \\ H_i &= h(K_i). \end{aligned}$$

and issues a smart card containing $\langle M_i, V_i, H_i, R_i, h(\cdot) \rangle$ to U_i.

Step 3. Upon receiving the smart card, U_i enters the random number b into its smart card. Accordingly, b is additionally stored on the smart card.

2.2 Mutual Authentication Phase

This phase constitutes the core of the scheme and is performed whenever some user U_i wants to log on to the server S_j. U_i initiates this phase by inserting its smart card into a card reader and entering its password PW_i and identity ID_i and the server S_j's identity ID_{S_j}. Given the user input, the smart card and the server perform the following steps as outlined in Fig. 2[1]:

[1] N_{rj} was incorrectly stated as N_{jr} in the thirty fifth line of in page 1121 of [2]. We have corrected this typographical error.

$$
\begin{array}{lcr}
\boxed{U_i}\ \langle PW_i \rangle & \boxed{S_j}\ \langle y \rangle & \boxed{RC}\ \langle x, y, r \rangle \\
K_i = & & \\
M_i \oplus h(ID_i || h(b \oplus PW_i)) & & \\
H_i' = h(K_i), H_i' \stackrel{?}{=} H_i & & \\
A_i = V_i \oplus h(b \oplus PW_i) & & \\
CID_i = & & \\
h(b \oplus PW_i) \oplus h(K_i || A_i || u) & & \\
P_i = K_i \oplus h(A_i || u || ID_{S_j}) & & \\
Q_i = h(V_i || A_i || u) & & \\
C_i = h(A_i || u + 1 || ID_{S_j}) & & \\
D_i = R_i \oplus ID_{S_j} \oplus u & & \\
\xrightarrow{\quad CID_i, P_i, Q_i, C_i, D_i, u \quad} & & \\
& L_j = h(ID_{S_j} || y) \oplus w & \\
& \xrightarrow{\quad L_j, ID_{S_j}, D_i, C_i, u \quad} & \\
& & w' = L_j \oplus h(ID_{S_j} || y) \\
& & R_i' = D_i \oplus ID_{S_j} \oplus u \\
& & Z_i' = R_i' \oplus h(x || r) \\
& & C_i \stackrel{?}{=} h(Z_i' || u + 1 || ID_{S_j}) \\
& & E_i = h(w' || h(ID_{S_j} || y) || n) \\
& & F_i = Z_i' \oplus h(h(ID_{S_j} || y) \oplus w') \\
& \xleftarrow{\quad E_i, F_i, n \quad} & \\
& E_i \stackrel{?}{=} h(w || h(ID_{S_j} || y) || n) & \\
& A_i' = F_i \oplus h(h(ID_{S_j} || y) \oplus w) & \\
& K_i' = P_i \oplus h(A_i' || u || ID_{S_j}) & \\
& \pi_i' = CID_i \oplus h(K_i' || A_i' || u) & \\
& V_i' = A_i' \oplus \pi_i' & \\
& Q_i \stackrel{?}{=} h(V_i' || A_i' || u) & \\
& G_i = h(V_i' || u || A_i' || ID_{S_j}) & \\
\xleftarrow{\quad G_i, t \quad} & & \\
G_i \stackrel{?}{=} h(V_i || u || A_i || ID_{S_j}) & & \\
J_i = h(V_i || t || A_i || ID_{S_j}) & & \\
\xrightarrow{\quad J_i \quad} & & \\
SK = h(V_i || A_i || u || t || ID_{S_j}) & SK = h(V_i || A_i || u || t || ID_{S_j}) &
\end{array}
$$

Fig. 2. The mutual authentication phase of Hsiang et al.'s scheme

Step 1. Using the user-given PW_i and ID_i, the smart card computes $K_i = M_i \oplus h(ID_i||h(b \oplus PW_i))$ and $H_i' = h(K_i)$. Then the smart card checks that H_i' is equal to H_i. If they are equal, the smart card proceeds to the next step. Otherwise, the smart card aborts the authentication phase.

Step 2. The smart card chooses a random nonce u and computes

$$\begin{aligned}
A_i &= V_i \oplus h(b \oplus PW_i),\\
CID_i &= h(b \oplus PW_i) \oplus h(K_i||A_i||u),\\
P_i &= K_i \oplus h(A_i||u||ID_{S_j}),\\
Q_i &= h(V_i||A_i||u),\\
C_i &= h(A_i||u+1||ID_{S_j}),\\
D_i &= R_i \oplus ID_{S_j} \oplus u.
\end{aligned}$$

Then the smart card sends the login request message $\langle CID_i, P_i, Q_i, C_i, D_i, u\rangle$ to the server S_j.

Step 3. S_j first chooses a random number w and computes $L_j = H(ID_{S_j}||y) \oplus w$. Then, S_j sends the message $\langle L_j, ID_{S_j}, D_i, C_i, u\rangle$ to the registration center RC.

Step 4. When the message arrives, the registration center RC computes $w' = L_j \oplus h(ID_{S_j}||y)$, $R_i' = D_i \oplus ID_{S_j} \oplus u$, and $Z_i' = R_i' \oplus h(x \oplus r)$. Then RC verifies that C_i equals $h(Z_i'||u+1||ID_{S_j})$. If this condition is untrue, RC rejects the login request. Otherwise, RC generates a new random nonce n, computes $E_i = h(w'||h(ID_{S_j}||y)||n)$, $F_i = Z_i' \oplus h(h(ID_{S_j}||y) \oplus w')$ and sends the response message $\langle E_i, F_i, n\rangle$ to the server S_j.

Step 5. After receiving $\langle E_i, F_i, n\rangle$, S_j chooses a random number t and computes

$$\begin{aligned}
A_i' &= F_i \oplus h(h(ID_{S_j}||y) \oplus w),\\
K_i' &= P_i \oplus h(A_i'||u||ID_{S_j}),\\
\pi_i' &= CID_i \oplus h(K_i'||A_i'||u),\\
V_i' &= A_i' \oplus \pi_i',\\
G_i &= h(V_i'||u||A_i'||ID_{S_j}).
\end{aligned}$$

Then, the server S_j checks that : (1) E_i is equal to $h(w||h(ID_{S_j}||y)||n)$ and (2) Q_i is equal to $h(V_i'||A_i'||u)$. If both of theses conditions hold, S_j believes U_i as authentic. Otherwise, S_j aborts its authentication attempt. After that, S_j sends the response message $\langle G_i, t\rangle$ to U_i

Step 6. Having received G_i from S_j, U_i verifies the correctness of G_i by checking that G_i equals $h(V_i||u||A_i||ID_{S_j})$. If correct, U_i accepts the response message, computes $J_i = h(V_i||t||A_i||ID_{S_j})$, and sends the message $\langle J_i\rangle$ to S_j; otherwise, reject it.

Step 7. After receiving $\langle J_i\rangle$, S_j verifies the correctness of J_i by checking that J_i equals $h(V_i'||t||A_i'||ID_{S_j})$. If correct, S_j believes as the legitimate user; otherwise, reject it.

Finally, U_i and S_j compute their session key SK as

$$SK = h(V_i||A_i||u||t||ID_{S_j}).$$

3 Cryptanalysis of Hsiang et al.'s Scheme

In this section we point out that Hsiang et al.'s scheme not only does not achieve its main security goal of authenticating between a remote individual and the server but also password security. In addition, we found that the scheme does not provide two-factor security.

3.1 Attacks against Mutual Authentication

Unfortunately, Hsiang et al.'s remote user authentication scheme supports neither user-to-server authentication nor server-to-user authentication. In this subsection, we show this by presenting two impersonation attacks, a user impersonation attack and a server impersonation attack. The user impersonation attack and the server impersonation attack are given to violate respectively user-to-server and server-to-user authentications of Hsiang et al.'s scheme.

Impersonating U_i to S_j. First, we present a user impersonation attack where an attacker U_a, who is a legitimate user registered with the registration center, can easily impersonate the remote user U_i to any other server S_j. Before describing the attack, we note that the secret values stored in a smart card could be extracted by monitoring its power consumption [4,8]. We now proceed to describe the user impersonation attack.

1. As a preliminary step, the attacker U_a extracts the secret values b, V_a and R_a stored in its smart card. From V_a and its hashed password value π_a, U_a recovers the value of $h(x \oplus r)$ as $h(x \oplus r) = V_a \oplus \pi_a \oplus R_a$.
2. Now when U_i initiates the authentication phase with the login request message $\langle CID_i, P_i, Q_i, C_i, D_i, u\rangle$, the attacker U_a posing as U_i intercepts this login request and sends to S_j a forged login message as follows: U_a who has obtained the login message, CID_i, P_i, u, D_i, and the secret value $h(x \oplus r)$, first, U_a chooses three random numbers, α, β, and $\mathtt{PW}'_i$, and computes

$$\begin{aligned}
R_i &= D_i \oplus ID_{S_j} \oplus u,\\
A_i &= R_i \oplus h(x \oplus r),\\
K_i &= P_i \oplus h(A_i||u||ID_{S_j}),\\
CID_i^* &= h(\beta \oplus \mathtt{PW}'_i) \oplus h(K_i||A_i||\alpha),\\
P_i^* &= K_i \oplus h(A_i||\alpha||ID_{S_j}),\\
V_i^* &= A_i \oplus h(\beta \oplus \mathtt{PW}'_i),\\
Q_i^* &= h(V_i^*||A_i||\alpha),\\
C_i^* &= h(Z_i||\alpha + 1||ID_{S_j}),\\
D_i^* &= R_i \oplus ID_{S_j} \oplus \alpha.
\end{aligned}$$

Then U_a posing as some registered user U_i sends the forged login request message $\langle CID_i^*, P_i^*, Q_i^*, C_i^*, D_i^*, \alpha\rangle$ as a login request message to the server S_j.

3. After receiving $\langle CID_i^*, P_i^*, Q_i^*, C_i^*, D_i^*, \alpha\rangle$, the server S_j chooses a random number w and computes $L_j = H(ID_{S_j}||y) \oplus w$. Then, S_j sends the message $\langle L_j, ID_{S_j}, D_i^*, C_i^*, \alpha\rangle$ to the registration center RC.
4. Since, from RC's point view, D_i^*, C_i^*, α are indistinguishable from D_i, C_i, u of an honest execution, RC believes that the message $\langle D_i^*, C_i^*, \alpha\rangle$ is from U_i. Hence, RC operates as specified in protocol using the received messages from S_j. First, RC computes w', R_i', Z_i' and proceeds to verify the authenticity of the login request. That is, RC computes $w' = L_j \oplus h(ID_{S_j}||y)$, $R_i' = D_i^* \oplus ID_{S_j} \oplus u$, and $Z_i' = R_i' \oplus h(x \oplus r)$ and checks that C_i^* equals $h((Z_i'||\alpha + 1||ID_{S_j})$. Since it hold, R$C$ will welcome U_a's visit to the system. Then, RC chooses a random number n, computes $E_i = h(w'||h(ID_{S_j}||y)||n)$ and $F_i = Z_i' \oplus h(h(ID_{S_j}||y) \oplus w)$ and sends the response message $\langle E_i F_i, n\rangle$ to S_j.
5. Since E_i, A_i', and Q_i are all valid, everything proceeds as usual. In response to U_a's login message, S_j chooses a random number t, computes $G_i = h(V_i'||u||A_i'||ID_{S_j})$, and sends $\langle G_i, t\rangle$ to U_a.
6. Now, an attacker U_a upon receiving G_i and t from S_j, computes $J_i = h(V_i||t||A_i||ID_{S_j})$, and sends the response message $\langle J_i\rangle$ to S_j. Because J_i equals $h(V'||t||A_i'||ID_{S_j})$, S_j will be unaware of attack and believes U_a as legitimate user U_i.

Impersonating S_j to U_i. Now, we present a server impersonation attack where an attacker U_a, who is a legitimate user registered with the registration center, can easily impersonate the remote server S_j to any other registered user U_i. Before describing the attack, we note that the secret values stored in a smart card could be extracted by monitoring its power consumption [4,8]. We now proceed to describe the server impersonation attack.

1. As a preliminary step, the attacker U_a extracts the secret values b, V_a and R_a stored in its smart card. From V_a and its hashed password value π_a, U_a recovers the value of $h(x \oplus r)$ as $h(x \oplus r) = V_a \oplus \pi_a \oplus R_a$.
2. When U_i initiates the authentication phase with the login request message $\langle CID_i, P_i, Q_i, C_i, D_i, u\rangle$, the attacker U_a posing as S_j intercepts this message and sends a forges the server S_j's response message as follows: U_a who has obtained the login message, CID_i, P_i, Q_i, D_i, u and the secret value, $h(x \oplus r)$, first
   (a) generates a random number γ.
   (b) computes
$$\begin{aligned} R_i &= D_i \oplus ID_{S_j} \oplus u, \\ A_i' &= R_i \oplus h(x \oplus r), \\ K_i' &= P_i \oplus h(A_i'||u||ID_{S_j}), \\ V_i' &= A_i' \oplus CID_i \oplus h(K_i'||A_i'||u), \\ G_a &= h(V_i'||u||A_i'||ID_{S_j}). \end{aligned}$$
   (c) and then sends $\langle G_a, \gamma\rangle$ in response to U_i's login request.
3. The forged response $\langle G_a, \gamma\rangle$ will pass the verification test by U_i since G_i is equal to $h(V_i||u||A_i||ID_{S_j})$. Hence, U_i believes U_a as the authentic server.

Security Analysis. It is clear that the vulnerability of Hsiang et al.'s scheme to two attacks above is mainly because the same hash value $h(x \oplus r)$ is used in computing Z_i for all users U_i: $Z_i = h(\pi_i||r) \oplus h(x \oplus r)$ for all i. This oversight allows the attacker to derive $h(x \oplus r)$ easily from Z_i by using the extracted value $R_i(= h(PW_i||r))$ from its smart card and then creates the vulnerability to two impersonation attacks. Based on this observation, anyone who knows the value of $h(x \oplus r)$ stored in its smart card is able to compute a valid login request $CID_i = h(b \oplus PW_i) \oplus h(K_i||A_i||u)$ (even without knowing U_i's password PW_i) or able to compute a valid response $G_i = h(V_i'||u||A_i'||ID_{S_j})$(even even without knowing the server's secret values x and r) .

3.2 Attack against Two-Factor Security

Loosely stated, two-factor security [12] means that the security of the scheme is provided when either the user's smart card or its password is stolen, but not both. Hsiang et al. [2] claim that their authentication scheme achieves two-factor security.

Assume that the attacker, who is a legitimate user registered with the registration center, has stolen the U_i's smart card or gained access to it and extracted the secret values stored in it by monitoring its power consumption [4,8]. Now the attacker U_a has obtained the value π_i stored in the U_i's smart card. Then the following description represents our off-line dictionary attack mounted by the attacker U_a against U_i's password: The attacker U_a, who wants to find out PW_i, now guesses possible passwords and checks them for correctness.

1. As a registered user, the attacker U_a, who has obtained R_a stored in its smart card, learns the value of $h(x \oplus r)$ by computing $h(x \oplus r) = V_a \oplus \pi_a \oplus R_a$ using its password PW_a.
2. Using V_i and R_i from U_i and $h(x \oplus r)$ from itself, U_a computes $h(b \oplus PW_i)$ $= V_i \oplus R_i \oplus h(x \oplus r)$.
3. Next, U_a makes a guess PW_i' for PW_i and computes $h(b \oplus PW_i')$.
4. U_a then verifies the correctness of PW_i' by checking the equality $h(b \oplus PW_i')$ $= h(b \oplus PW_i)$. Notice that if PW_i' and PW_i are equal, then $h(b \oplus PW_i') = h(b \oplus PW_i)$ ought to be satisfied.
5. U_a repeats steps (3) and (4) until a correct password is found.

The two above security vulnerabilities of Hsiang et al.'s scheme are attributed to the following a flaw in its design:

- The same hash value $h(x \oplus r)$ is used in computing Z_i for all users U_i: $Z_i = h(\pi_i||r) \oplus h(x \oplus r))$ for all i.

4 Conclusion

We have analyzed the security of the smart card based user authentication scheme proposed by Hsiang et al. [2]. Our security analysis uncovered that Hsiang et al.'s scheme does not achieve its main security goal not only of authenticating between a remote individual and the server but also of password security. Now, the failure of Hsiang et al.'s scheme to achieve authentication has been made clear through two attacks, a server impersonation attack and a user impersonation attack, on the scheme. The server impersonation attack and the user impersonation attack have been considered to infringe respectively server-to-user and user-to-server authentications of the scheme. In addition, we demonstrate that Hsiang et al.'s scheme fails to provide two-factor security which guarantees the security of the scheme when either the user's smart card or its password is stolen, but not both by employing the off-line dictionary attack.

References

1. Chang, C., Lee, J.S.: An efficient and secure multi-server password authentication scheme using smart cards. In: Proceeding of the International Conference on Cyberworlds, IEEE, Los Alamitos (2004)
2. Hsiang, H.-C., Shih, W.-K.: Improvement of the secure dynamic Id based remote user authentication scheme for multi-server environment. Computer Standards & Interfaces 31, 1118–1123 (2009)
3. Juang, W.S.: Efficient multi-server password authenticated key agreement using smart cards. IEEE Trans. Consum. Electron. 50(1), 251–255 (2004)
4. Kocher, P.C., Jaffe, J., Jun, B.: Differential Power Analysis. In: Wiener, M. (ed.) CRYPTO 1999. LNCS, vol. 1666, pp. 388–397. Springer, Heidelberg (1999)
5. Li, L.-H., Lin, I.-C., Hwang, M.-S.: A remote password authentication scheme for multi-server architecture using neural networks. IEEE Trans. on Neural Networks 12(6), 1498–1504 (2001)
6. Liao, Y.-P., Wang, S.-S.: A secure dynamic ID based remote user authentication scheme for multi-server environment, Computer Standards & Interfaces (2007) (article in press)
7. Lin, I.-C., Hwang, M.-S., Li, L.-H.: A new remote user authentication scheme for multi-server internet environments. Future Generation Computer System 19, 13–22 (2003)
8. Messergers, T.S., Dabbish, E.A., Sloan, R.H.: Examining smart card security under the threat of power analysis attacks. IEEE Trans. Comput. 51(5), 541–552 (2002)
9. Tsuar, W.-J.: An enhanced user authentication scheme for multi-server internet services. Appl. Math. Comput. 170, 258–266 (2005)
10. Tsuar, W.-J., Wu, C.-C., Lee, W.-B.: A flexible user authentication scheme for multi-server internet services. In: Lorenz, P. (ed.) ICN 2001. LNCS, vol. 2093, pp. 174–183. Springer, Heidelberg (2001)
11. Tsaur, W.-J., Wu, C.-C., Lee, W.-B.: A smart card-based remote scheme for password authentication in multi-server Internet services. Computer Standards & break Interfaces 27, 39–51 (2004)
12. Tian, X., Zhu, R.W., Wong, D.S.: Improved efficient remote user authentication schemes. Int. J. Netw. Secur. 4(2), 149–154 (2007)
13. Yang, Y., Wang, S., Bao, F., Wang, J., Deng, R.: New efficient user identification and key distribution scheme providing enhanced security. Comput. Secur. 23(8), 697–704 (2004)

Exploring Informational Privacy Perceptions in the Context of Online Social Networks: A Phenomenology Perspective

Emma Nuraihan Mior Ibrahim

Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, 40450
Shah Alam, Selangor, Malaysia
emma@tmsk.uitm.edu.my

Abstract. The paper presents a conceptual framework of informational privacy dimensions and its elements in the context of social networks that derived from the understanding of the user's perceptions based on a focus group interview. Phenomenology approach is adopted as it enables us to discover the lived experiences of the people involved rather using a survey or a self reported attitude studies for online social networks investigation where users tend to depart from their usual practices which is proven to be biased. It is found that the main dimensions of the informational privacy within the context of online social networks are (1) limited communication dimension that has the element of sharing of personal information with trusted others (2) release of personal information dimension that has the element of content control and the amount of personal information being released; and (3) control dimension that has the element of the ability to control over information about oneself.

Keywords: informational privacy, online social networks, phenomenology method.

1 Introduction

For the past several years, technology particularly the advent of Internet has played major role in changing the lifestyles of Internet users worldwide where it is proven to be the most vital online communication medium for the distribution, retrieval and sharing of information since the mid 90's [1]. Many activities can be done at home, workplace or anywhere in the world including communicating through online social networks. Nevertheless, the Internet was invented as an "insecure communications" media that lead to several issues including informational privacy [2]. Thus, this research will look into the new phenomenon of explosive growth of online social networks as it has gain interests from researchers and businesses worldwide [3] regarding the possibilities and the risks. Furthermore, the most controversial issues arise within online social networks pertaining to their implications and impacts on

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 330–338, 2011.

informational privacy issues as research has proven that users are more willing to disclose their personal information to public without much concern towards their informational privacy states [4]. In addition, due to the informational privacy issues such as media-driven user informational privacy concerns, the continuous informational privacy-related public scandals and media reporting the questionable information handling practices has began to boost user awareness regarding the informational privacy threats they encounters [5]. However, there is yet systematic research on informational privacy that can cater the needs of the users on the protection of informational privacy specific to online social networks context This problem may arise due to the lack of suitable measurement instrument of informational privacy confining the context of online social networks solely [6]. Thus, the research attempts to look into the user's understanding of informational privacy from the phenomenology perspective based on their real experience communicating through the online social networks in order to find possible dimensions and elements of informational privacy.

2 Research Background

2.1 Concept of Privacy

Privacy is a common term that is used widely in every day's life which is perceived in different meaning due to its multifaceted concepts and dimensions that depend on the value and scope in various domains such as ecommerce, businesses, Internet banking and society [6] [7] [8] [9] [10]. However in the literature there are several dimensions that have been discussed. For example, Burgoon et. al [8] identified four dimensions of privacy which they defined privacy as "the ability to control and limit physical, interactional psychological and informational access to the self or one's group". Meanwhile, Fischer- Hübner [9] have outlined three dimensions of privacy consisting personal physical privacy, territorial privacy and informational privacy which is much similar to DeCew [10] that classified privacy into three dimensions namely informational, accessibility, and expressive privacy [6] [10]. However, these dimensions are overlap with each other (see Table 1).

Definition of Informational Privacy. The term informational privacy, is frequently used whilst referring to individuals' online privacy. Westin [11] had described informational privacy as the claim of individuals, groups or institutions to determine for themselves when, how and to what extent information about them is communicated to others. This definition has been used as a guideline for the formulation of fair information practices (FIP), which is a set of policy guidelines that is used worldwide [12]. Meanwhile, Internet Society [13] defined informational privacy as "the right of an entity (normally a person), acting in its own behalf, to determine the degree to which it will interact with its environment, including the degree to which the entity is willing to share information about itself with others".

Table 1. Summary of Privacy Dimensions from the Literature

Dimensions	Definitions
Informational Privacy	-Personal information that includes finances and medical details in which how much an individual's are willing to share this personal information to another person or company and to what extent the information will be disclosed [9] [11].
Bodily Privacy	-Also known as 'accessibility' or physical privacy [8] [9] [10]. The physical access of an individual to others whereby an individual will determine who can have body contact and share their sense of perception and observation with them.
Privacy of Communications	-The security and privacy of mail, telephone calls, email and other forms of communication [8].
Territorial Privacy	-Also known as 'expressive' or interactional privacy [10]. It relates to protection of autonomy for an individual to express their self-identity throughout speech or activity without being violated from interference, pressure and coercion from government or from other persons. Therefore, territorial privacy improved individuals self expression and capability to build interpersonal relationships but restricted the social control over lifestyle choices [9].

2.2 Online Social Networks

Online social networks can be described as online environment where people generate public or semi-public self-descriptive profiles as well as building connections to other people [14]. Individuals can produce impression on others as online social networks stress on the oral and linguistic signs [15]. This feature allows users to expose their identity fully plus their hidden identities [16] as they can reports their experiences and achievements, upload photos, and display their list of connections, that shows their pleasing connections and intuitions. In addition, online social networks help developing user's identity through the application usage, board discussions, blogging, or posting commentaries in public areas [5]. Thus, online social networks give the users opportunities to pre-decide the information they want to publish to public that portray their identity closer to their "ideal self" as proven by the study of Boyd [14], he claimed that users of online social networks are more likely to show their prominent identity to impress significant others by exposing their true identity but this is limited to connection with real-life friends differ from anonymous online forums, dating platforms, and chats where users tend to exaggerate their true identity and self achievement [17].

2.3 Informational Privacy in the Online Social Networks Environment

In the context of online behavior, informational privacy is at risk. Studies noted that even though protecting individuals personal information still remain questionable, most individuals spend time online based on both social and non-social activities whilst submitting personal information online for instance, building social networking on social networking sites. To maximize the benefits of online social networks users have to bring up the best impression about themselves to public as it is part of online social networks activities. Nevertheless, online social networks bring informational privacy concern to users as users are more cautious to do online social networks activities [5] due to informational privacy threats [14]. As a result, users begun to change the way they disclose their personal information in online social networks as they have lack of control on the information. This is a concerned situation as this changing might harm the sustainability of online social networks since users are incapable to construct their identity in the preferred way in which they might diminish their participation in online social networks or even worst leaving the network. This may cause the loss of commercial and public value of the online social networks platforms [5]. Thus, there is a need to understand the potential of online social networks where user's informational privacy concern should be taken into account. Even though, several researches in the past has studied informational privacy within the online social networks [18] [19] [20] but the results are insufficient and often of debatable in nature [5]. There are still few systematic research has arrived with a conceptual framework on the dimensions and elements of informational privacy from the user's understanding. This is significant as individual's understanding of informational privacy might be different from others based on their own personal characteristics and past experiences that have cultural influences [21]. Furthermore, individuals might have different opinions on what is fair and what is not fair regarding the collection and use of their information by the online social networks.

3 Problem Background

The area of informational privacy in online social networks is being strictly under researched [22] as most literature on online social networks mainly focuses on the issues of information disclosure, management of informational privacy settings and informational privacy awareness [19] [22] [23]. Besides, various studies on the impact of informational privacy concerns on behavior [18] [19] [20] are insufficient and often of debatable in nature as the studies are inconsistent in terms of the results due to the lack of suitable instrument that measure informational privacy concerns within the online social networks solely [5]. Thus, this research focuses on the right as individuals which is the users' of online social networks as there is limited number of studies that focus on user informational privacy concerns within the online social networks. In addition, much of the published literature on informational privacy in the past only concentrates on protecting some types of data regardless concerning the perception of users on what considered as private data. However, these expert opinion on the issues of informational privacy might not be sufficient in determining a good communication technology or any effective policies to use in which it do not reflect the needs of users making it vital to explore the understandings of users on what will be considered invasive to them and the reasons. Meanwhile, in terms of methodology, the methods used in exploring informational privacy will usually used surveys [24]. Moreover, Harper and Singleton [25] claimed that in a review of 23 privacy surveys, "surveys may suffer from manipulative questioning on one hand, and that unprompted surveys may reveal very little privacy concern on the other hand" making it researcher-biased because sometimes it is hard for researcher to recognize certain informational privacy issues [24].

4 Research Design and Approach

4.1 Phenomenology Method

Phenomenology method is used in exploring the understanding on the notion of informational privacy among online social networks users that focuses on personal consciousness where it study on the real ‘phenomena’ aiming at the root of the thing itself [26] [27]. The research aim is to describe the phenomenon as accurate as possible, abstaining from any pre-given framework, but consistent to the facts. In addition, the study deals with the lived experiences, the user’s experiences as the first person point of view of the people involved, or who were involved, with the issue that is being researched. The researcher listened to the recording and take notes to explicit participants key words, phrases and statements. Field note can be consider as part of data analysis since it involve interpretation thus the researcher must remember not to be bias while doing field note as the root of phenomenology that deals with the lived experiences of the participants in the research. However, there is no specific guideline to conduct phenomenology research [28]. It is pointed out that researchers who use phenomenology is hesitant to set techniques corresponded with. Thus the research design is adapted from Moustakas [27].

4.2 Focus Group Interview

Due to the exploratory nature of the research, the empirical study is conducted via a semi structured interview session with six participants [29] who have at least 2 years experiences in conducting online transactions and engaging in other social networking sites such as Twitter, MySpace, Friendster, Facebook actively on their understandings of the informational privacy within the online social networks environment. These participants are more likely to have view on offline and online risks where they commonly participating with the Internet activities [30]. Furthermore, the participants chosen are having at least bachelor degree qualification as individuals from educated background will usually have some experiences using technology [31]. In addition, the preparation for the interview questionnaire are designed and discussed from synthesizing privacy and informational privacy literature including theories and its dimensions based on the work of [4] [5] [11] [31] [32] [33]. However, these literatures are only to form the body of arguments to assist the researcher in structuring the interview questionnaire and not the pre-determined set of framework. Nevertheless, the researcher remains unbiased and letting the participants elaborate their perceptions and experiences in their own terms.

5 Results and Discussions

The general and unique themes from the entire interview is extorted to make a merged summary in order to identify the most concerned dimensions and elements of informational privacy based on the user’s understanding in the context of online social networks as shown in Table 2 below.

Table 2. Summary of the results on the informational privacy dimensions

Themes	Meanings
Solitude Seclusion /Isolation	Control of the "interactional borders "or privacy boundary
Intimacy	Control over one's intimate relationships
Reserve	Limit knowledge of one Control one's personal information Limitation of disclosure to others Boundary of exposure of individuals' personal information to others
Limited Communication	Sharing of personal information with trusted others
Secrecy	Disguise of certain issues from others
Control/ Personhood	Humiliation Manipulation of personal information that involve one's personality Manipulation of personal information that can cause humiliation and discrimination
Control over Personal Information about Oneself	Ability to control over information about oneself
Release of Personal Information	Control the content and amount of personal information being released
Distribution of Personal Information	Right to distribute own personal information
Emotional Release	Relieve, shared emotions
Self-evaluation	Individuals personal experiences that help them to make decision
Usage of information/Use of Personal Information	Purposeful usage of information
Information Compensation	Information trade-off for benefits
Identity Theft	Manipulation of personal information that can lead to fraud
Security	Security in disclosing of personal information
Inappropriate/Unrelated	Inappropriate/Unrelated association with oneself
Options	Options on what kind of information revelation is needed
Misuse of information	Misuse of information to other unknown party

From the results of the interview, it is found that there are some new themes derived from the user's perceptions which are *security, inappropriate/ unrelated association with oneself, options and misuse of information*. However, the other nine themes that is significant to the users are similar with the existing themes from the literature that are *solitude, intimacy, reserve*, also known in the literature as limited communication and secrecy; *control* which is also known as personhood, control over personal information about oneself*, release of personal information* and *distribution of personal information;, emotional release, self-evaluation, usage of information/use of personal information, information compensation and identity theft.*

Nevertheless it can be summarized that informational privacy in the context of online social networks is reflected by some significant dimensions that is, *limited communication* in which it is defined as the sharing of information to certain people. Users only shared their personal information with certain people, the one that they trusted. This is because according to the participants, they value the informational privacy and do wished that their informational privacy is protected from unauthorized access by third parties. Furthermore, informational privacy has also been viewed as the r*elease of personal information dimension* which is the control of the content and amount of personal information being released to others. Participants argued that the online social networks are a medium of information sharing, where they believed that if someone wants to make information private they do not have to share and put the information on their online social network account. This is due to the fact that the main intention of joining online social networks itself is to build connections to other people that may lead to the disclosed of personal information.

In addition, c*ontrol* has been reflected as informational privacy in which control is defined as the ability to control over information about oneself. Participants believed that informational privacy allows users to choose whether they want to share or not to share their personal information as it lets users to control their own private information and the access of the information to certain people. Thus, users of online social networks may think that they lose their privacy if they cannot decide who can collect, use, share their personal information in other word, the control of the personal information collected.

Furthermore, the duality of the phenomenon of informational privacy in the context of online social networks occurs whereby the more the users are exposed their networking to the social medium, the limited the personal information that the users are willing to disclose. This can be explained by the proposed framework of informational privacy within the online social networks as shown in Figure 1 below whereby it can be seen that the interplay of release of personal information, limited communication, and control construct. This can be reflected in a situation in which even though it seems that users of online social networks is disclosing their personal information to public in order to build connections to other people the fact is, these users are actually control of the content and amount of personal information being released to others which they only shared the personal information with the trusted others. In addition, both of release of personal information constructs and limited communication construct is influence by the control construct in the sense that controlling over information about oneself is needed as users of online social networks feel that their privacy is invaded if they have no control on their own personal information.

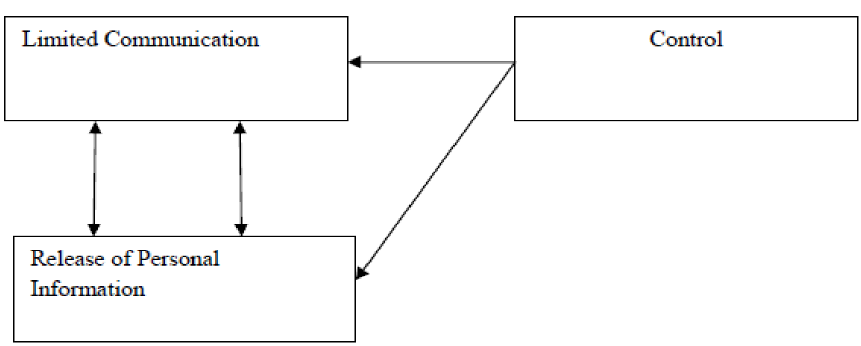


Fig. 1. The conceptual framework of informational privacy in the online social networks

6 Conclusion and Future Directions

The study have found four new emerging themes informational privacy in the context of online social networks that is s*ecurity*, *inappropriate/unrelated association with oneself,* o*ption and* m*isuse of information* which is not significantly addressed previously in the existing literature. These themes will contribute to the existing dimensions in the informational privacy literature and need further investigation. However, the findings of the research are based on a small sample group made up of only six participants as this research is an exploratory in nature that only tends to find the possible constructs of informational privacy dimensions and its elements. Thus, it is recommended to carry out a similar research in a quantitative manner involving larger respondents to examine the relationship between the proposed constructs of the conceptual framework.

References

1. Boncle., Good privacy. Networking that embraces privacy (2008), http://www.boncle.com/index.html (retrieved February 3, 2010)
2. Givens., Online Privacy and Technology. Privacy Rights Clearing House (2009), http://www.privacyrights.org/Online-Privacyand-Technology (retrieved February 3, 2010)
3. Caudill, E.M., Murphey, P.E.: Consumer Online Privacy: Legal and Ethical Issues. Journal of Public Policy and Marketing 19(1), 7–19 (2000)
4. Sheehan, K.B., Hoy, M.G.: Dimensions of privacy concern among online consumers. Journal of Public Policy & Marketing 19(1), 62–73 (2000)
5. Krasnova, H., Günther, O., Spiekermann, S., Koroleva, K.: Privacy concerns and identity in online social networks. Identity Journal Limited 2009 (2009)
6. Schofield, C., Joinson, A.D.: Privacy, Trust, and Disclosure Online (2008)
7. Stern, T.: Internet Privacy from the Individual and Business Perspectives – Introduction: privacy, Trust: privacy from the business perspective, Individual vs. business = privacy vs. personalization, Cookies, Google, Solutions. Baruch College, CUNY, USA online encyclopaedia 2009 (2009)
8. Burgoon, J.: Privacy and communication. In: Burgoon, M. (ed.) Comms Yearbook 6, pp. 206–249. Sage, Forrester Research, CA, Cambridge, MA (1982)
9. Fischer-Hübner, S.:IT-Security and Privacy-Design and Use of Privacy-Enhancing Security Mechanisms. LNCS, vol. 1958. Springer, Heidelberg (2001)
10. DeCew, J.: In pursuit of privacy: Law, ethics, and the rise of technology. Cornell University Press, Ithaca (1997)
11. Westin, A.: Privacy and freedom. Atheneum, New York (1967)
12. Lederer, S., Hong, J.I., Dey, A.K., Landay, J.A.: Personal privacy through understanding and action: five pitfalls for designers. Personal Ubiquitous Computing 8(6), 440–454 (2004)
13. Internet Society (2007), http://www.faqs.org/rfcs/rfc2828.html (retrieved February 3, 2010)
14. Boyd, D., Ellison, N.: Social Network Sites: Definition, History, and Scholarship. Journal of Computer-Mediated Communication 13(1), 11 (2007)
15. Ellison, N., Heino, R., Gibbs, J.: Managing impressions online: Self-presentation processes in the online dating environment. Journal of Computer-Mediated Communication 11, 2 (2006), article 2

16. Schau, H.J., Gilly, M.C.: We are what we post? Self-presentation in personal web space. Journal of Consumer Research 30(3), 385–404 (2003)
17. Brym, R.J., Lenton, R.L.: Love Online: A Report on Digital Dating in Canada (2001), http://www.nelson.com/nelson/harcourt/sociology/newsociety3e/loveonline.pdf (2008) (retrieved March 10, 2010)
18. Dwyer, C., Hiltz, S.R., Passerini, K.: Trust and Privacy Concern within Social Networking Sites: A Comparison of Facebook and MySpace. In: Thirteenth American Conference on Information Systems, USA (2007)
19. Tufekci, Z.: Can you see me now? Audience and disclosure regulation in Online Social Network Sites. Bulletin of Science, Technology & Society 28(1), 20–36 (2008)
20. Stutzman, F.: An Evaluation of Identity-Sharing Behavior in Social Network Communities. International Digital and Media Arts Journal 3(1), 10–13 (2006)
21. Moor, J.: Towards a theory of privacy in the information age. Computers and Society 27(3), 27–32 (1997)
22. Preibusch, S., Hoser, B., Gürses, S., Berendt, B.: Ubiquitous social networks opportunities and challenges for privacy-aware user modelling. In: Proceedings of the Data Mining for User Modelling Workshop (DM.UM 2007) at UM 2007, Corfu (June 2007)
23. Krishnamurthy, B., Wills, C.E.: Characterizing privacy in online social networks. In: Proceedings of the First Workshop on online Social Networks, WOSP 2008, Seattle, WA, USA, August 18, pp. 37–42. ACM, New York (2008)
24. Patil, S., Romero, N., Karat, J.: Privacy and HCI: Methodologies for Studying Privacy Issues. In: CHI 2006, Montréal, Québec, Canada, April 22-27 (2006)
25. Harper, J., Singleton, S.: With a Grain of Salt: What Consumer Privacy Surveys Don't Tell Us (2001), http://ssrn.com/abstract=299930 (retrieved February 3, 2010)
26. Kruger, D.: An introduction to phenomenological psychology, 2nd edn., Cape Town (1988)
27. Moustakas, C.: Phenomenological research methods. Sage, Thousand Oaks (1994)
28. Groenewald, T.: A phenomenological research design illustrated. International 140 Journal of Qualitative Methods 3(1) (2004), Article 4
29. Cairns, P., Cox, A.L.: Research Methods for Human Computer Interaction, Cambridge (2009)
30. Nyshadham, E.A., Ugbaja, M.: A study of Ecommerce Risk Perceptions Among B2C Consumers: A Two Country Study. In: Proceedings of the 19th Bled eConference, Bled, Slovenia, June 5-7 (2006)
31. Malhotra, N.K., Kim, S.S., Agarwal, J.: Internet users' information privacy concerns (IUIPC): The construct, the scale and a causal model. Information Systems Research 15, 336–355 (2004)
32. Taylor, D.G., Davis, D.F., Jillapalli, R.: Privacy concern and online personalization: the moderating effects of information control and compensation. Electron Commer. Res. 9, 203–223 (2009)
33. Carew, P.J., Stapleton, L.: Towards a Privacy Framework for Information Systems Development. Information Systems Development: Advances in Theory, Practice and Education (2005)

Server-Aided Password-Authenticated Key Exchange: From 3-Party to Group*

Junghyun Nam[1], Juryon Paik[2], Jeeyeon Kim[2],
Youngsook Lee[3], and Dongho Won[2,**]

[1] Department of Computer Engineering, Konkuk University, Korea
jhnam@kku.ac.kr
[2] Department of Computer Engineering, Sungkyunkwan University, Korea
wise96@ece.skku.ac.kr,jeeyeonkim@paran.com,dhwon@security.re.kr
[3] Department of Cyber Investigation Police, Howon University, Korea
ysooklee@howon.ac.kr

Abstract. Protocols for group key exchange are cryptographic algorithms that describe how a group of parties communicating over a public network can come up with a common secret key. Due to their critical role in building secure multicast channels, a number of group key exchange protocols have been proposed over the years for a variety of settings. In this work, we present a new protocol for password-authenticated group key exchange in the model where the clients wishing to establish a common secret do not share any password between them but hold their individual password shared with a trusted server. This model is practical in that no matter how many different session keys for different groups a client wants to generate, he/she does not need to hold multiple passwords but only needs to remember a single password shared with the server. Our construction is generic. We assume a 3-party password-authenticated key exchange protocol and use it as a key component in building our password-authenticated GKE protocol. Our generic protocol requires no further long-term secrets than those used in the underlying 3-party protocol. This implies that if the given 3-party protocol is password-only authenticated, then our group key exchange protocol is password-only authenticated as well.

Keywords: Group key exchange, multicast, 3-party key exchange, password.

1 Introduction

The increasing ubiquity of computer networks is accelerating the development of group-oriented applications in which a group of parties communicate collaboratively to achieve their common interest or objective. Typical group-oriented applications include video/audio teleconferencing, distributed multiplayer games,

* This work was supported by Priority Research Centers Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2010-0020210).

** Corresponding author.

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 339–348, 2011.

grid computing, collaborative workspaces, and social networking services. In particular, social networking services such as Twitter [25] and Facebook [12] have recently gained tremendous popularity and are redefining our sense of community. The proliferation of group-oriented applications has led to a growing concern in security of group communications. The current Internet, by design, is an open network which might be controlled by an adversary. Today's adversaries are equipped with more powerful computing resources and attacking tools than ever before.

One valuable tool for protecting group communications is protocols for group key exchange (GKE). A group of parties communicating over a public network can generate a common secret key (called a *session key*) by running a GKE protocol. Once a session key has been established, the parties can use this key to encrypt and/or authenticate their subsequent multicast messages. This represents a typical way of communicating confidentially and with integrity over a public channel. The session key, of course, must be known only to the intended parties at the end of the protocol run, because otherwise the whole system becomes vulnerable to all manner of attacks. Roughly stated, a key exchange protocol satisfying this requirement is said to be *authenticated*. Due to their significance in building secure multicast channels, authenticated GKE protocols have been extensively studied over the last decades [13,7,8,20,17,6,18,5,15,4,16,14,2,21,1,27]. However, despite all the research efforts made so far, the design of an authenticated GKE protocol is still notoriously hard. Many GKE protocols, even some with a claimed proof of security, have been analyzed to be vulnerable to a certain kind of attack years after they were published [23,11,24,22].

Any protocol for authenticated key exchange inherently requires that the protocol participants establish their long-term authentication secrets (either low-entropy passwords or high-entropy cryptographic keys) before they ever run the protocol. Protocols for password-authenticated key exchange are designed to work even when the authentication secrets are human-memorable passwords chosen from a small known set of values. These password-based protocols, despite their practical significance in today's computing environments, are notoriously hard to design right. The major hurdle to password-authenticated key exchange is (off-line) dictionary attacks in which an adversary exhaustively enumerates all possible passwords in an off-line manner to find out the correct one. In this work, we present a new protocol for password-authenticated GKE. Like the previous protocols of [9,10,19,27], Our protocol assumes a Kerberos-like authentication model in which the clients trying to establish a common secret do not share any password between them but hold their individual password shared with a trusted server. The role of the trusted server in this model is to provide the clients with a centralized authentication service. This model enjoys the obvious practical advantage that no matter how many different session keys for different groups a client wants to generate, he/she does not need to hold multiple passwords but only needs to remember a single password shared with the server. Our protocol differs from previous designs [9,10,19] in that it is constructed generically from a password-authenticated 3-party key exchange protocol secure against active adversaries. Although Yi et al.'s protocol [27] features generic construction as well, it is different from ours in two aspects: (1) it can be constructed from a GKE protocol secure against passive adversaries and (2) it employs identity-based cryptography where an arbitrary identity like an email address can serve as a public key. Hence, the best way to describe our protocol is the first server-aided password-authenticated GKE protocol that builds on a password-authenticated 3-party key exchange protocol.

In the next section, we give some preliminaries required for the security proof of the proposed protocol, including a communication and adversary model with an associated definition of security. Then, in Section 3, we present our protocol for server-aided password-authenticated GKE protocol and consider its efficiency and security.

2 Formal Setting

Any form of security analysis of a cryptographic construction should be preceded by clear definitions of its security goals and tools. In this section we provide such a preliminary formalism for server-aided password-authenticated GKE.

2.1 Communication and Adversary Model

Participants/Passwords. There are two types of participants in a server-aided password-authenticated GKE protocol: clients and servers. Let $\mathcal{C}$ be the set of all clients and $\mathcal{S}$ be the set of all servers. $\mathcal{C}$ and $\mathcal{S}$ are assumed to be of polynomial size. Before the protocol is executed for the first time, each client $C_i \in \mathcal{C}$ chooses a password pw_i from a dictionary D of size d and then registers it on a server $S \in \mathcal{S}$. The clients in any subset of $\mathcal{C}$ may run the protocol with a server $S \in \mathcal{S}$ to establish a common key, as long as all the clients (in the subset) have registered their individual passwords with the same server S. Each client may run the protocol multiple times either serially or concurrently, with possibly different sets of participants. Thus, at a given time, there could be many instances of a single client. We use C_i^π to denote the π-th instance of client C_i. All instances of a client C_i use the same password pw_i even if they participate in their respective sessions independently.

Partners. Intuitively, the *partners* of an instance is the set of all instances that should compute the same session key as the instance in an execution of the protocol. Like most of previous works, we use the notion of *session IDs* to define partnership between instances. Literally, a session ID (denoted as sid) is a unique identifier of a communication session. In our protocol, session IDs are constructed during protocol runs. We also need the notion of *group IDs* to define partnership properly. A group ID (denoted as gid) is a set consisting of the identities of the protocol participants. This notion is clearly natural because it is impossible (not even defined) to ever execute a group key exchange protocol without participants. Indeed, a group ID is a both necessary and important input to any protocol execution. We use sid_i^π and gid_i^π to denote respectively sid

and gid of instance C_i^π. Note that gid_i^π includes C_i itself and its trusted server S. Session IDs and group IDs are public and assumed to be available to the adversary.

An instance is said to *accept* when it successfully computes a session key in a protocol execution. Let acc_i^π be a boolean variable that evaluates to TRUE if C_i^π has accepted, and FALSE otherwise. We say that any two instances C_i^π and C_j^ω are *partners* of each other, or equivalently, *partnered* iff all the following three conditions are satisfied: (1) $\mathsf{sid}_i^\pi = \mathsf{sid}_j^\omega$, (2) $\mathsf{gid}_i^\pi = \mathsf{gid}_j^\omega$, and (3) $\mathsf{acc}_i^\pi = \mathsf{acc}_j^\omega =$ TRUE. We also say that two instances C_i^π and C_j^ω are *potential partners* of each other, or equivalently, *potentially partnered* iff the first two conditions above hold. We use pid_i^π and ppid_i^π to denote respectively the partners and the potential partners of the instance C_i^π. Then it follows by the definitions that $\mathsf{pid}_i^\pi \subseteq \mathsf{ppid}_i^\pi$.

Adversary. The adversary in our model controls all message flows of the protocol and can ask participants to open up access to any long-term secrets and session keys. These capabilities of the adversary are modeled via various oracles to which the adversary is allowed to make queries.

– $\mathsf{Execute}(\mathsf{gid})$: This query prompts the clients and the server in gid (precisely, their instances) to execute the protocol. The transcript of the execution is returned to the adversary as the output of the query. This models passive attacks on the protocol.
– $\mathsf{Send}(U^\pi, M)$: Let U^π be an instance of either a client or a server. This query sends message M to instance U^π. The instance U^π proceeds as it would in the protocol upon receiving message M; the instance updates its state performing any required computation, and generates and sends out a response message as needed. The response message, if any, is the output of this query and is returned to the adversary. This models active attacks on the protocol, allowing the adversary to control at will all message flows between instances. A query of the form $\mathsf{Send}(U^\pi,$ "`start`":$\mathsf{gid})$ prompts U^π to initiate an execution of the protocol in which the participants are the clients and the server specified in gid.
– $\mathsf{Reveal}(C_i^\pi)$: This query returns to the adversary the session key held by C_i^π. This oracle call captures the idea that exposure of some session keys should not affect the security of other session keys. The adversary is not allowed to ask this query if it has already queried $\mathsf{Test}(C_j^\omega)$ for some $C_j^\omega \in \mathsf{pid}_i^\pi$ (see below for the description of the Test oracle).
– $\mathsf{Corrupt}(U)$: Let U be a client or a server. This query returns to the adversary all long-term secrets of U. This models the adversary's capability of breaking into U's machine and gaining access to the long-term data set stored there. The adversary can issue this query at any time regardless of whether U is currently executing the protocol or not. This oracle call captures the idea that damage due to loss of U's long-term secrets should be restricted to those sessions where U will participate in the future.
– $\mathsf{Test}(C_i^\pi)$: This query provides a means of defining security. The output of this query depends on the hidden bit b that the Test oracle chooses uniformly

at random from $\{0, 1\}$ during its initialization phase. The Test oracle returns the real session key held by C_i^π if $b = 1$, or returns a random session key drawn from the key space if $b = 0$. The query can be asked only when instance C_i^π is *fresh* (see Section 2.2 for the definition of freshness). The adversary is allowed a single Test query, at any time during its execution.

Definition 1. *An adversary is called* active *iff it is allowed to access all the oracles described above, and called* passive *iff it is allowed to access all but the* Send *oracle.*

We represent the amount of queries used by an adversary as an ordered sequence of five non-negative integers, $Q = (q_{\text{exec}}, q_{\text{send}}, q_{\text{reve}}, q_{\text{corr}}, q_{\text{test}})$, where the five elements refer to the numbers of queries that the adversary made respectively to its $\mathsf{Execute}$, Send, Reveal, $\mathsf{Corrupt}$, and Test oracles. We call this usage of queries by an adversary the *query complexity* of the adversary. Note that by Definition 1, the query complexity of a passive adversary is always represented as a sequence of the form $Q = (q_{\text{exec}}, 0, q_{\text{reve}}, q_{\text{corr}}, q_{\text{test}})$.

2.2 Security Definition and Assumptions

Freshness. The notion of *freshness* is used in the definition of security to prohibit the adversary from asking the Test query against an instance whose session key (or some information about the key) can be exposed by trivial means.

Definition 2. *The instance C_i^π is considered unfresh iff any of the following conditions hold:*

1. $\mathsf{acc}_i^\pi = \text{FALSE}$.
2. *The adversary queried* $\mathsf{Corrupt}(U)$ *for some* $U \in \mathsf{gid}_i^\pi$ *before some instance in* ppid_i^π *accepts.*
3. *The adversary queried* $\mathsf{Reveal}(C_j^\omega)$ *for some* $C_j^\omega \in \mathsf{pid}_i^\pi$.

All other instances are considered fresh.

Security. The security of a server-aided password-authenticated GKE protocol P against an adversary $\mathcal{A}$ is defined in terms of the probability that $\mathcal{A}$ succeeds in distinguishing random session keys from real session keys established by the protocol P. That is, the adversary $\mathcal{A}$ is considered successful in attacking P if it breaks the semantic security of session keys generated by P. This notion of security is defined in the context of the following two-stage game, where the goal of adversary $\mathcal{A}$ is to correctly guess the value of the hidden bit b chosen by the Test oracle.

- **Stage 1:** $\mathcal{A}$ makes any allowed oracle queries at will as many times as it wishes.
- **Stage 2:** Once $\mathcal{A}$ decides that Stage 1 is over, it outputs a bit b' as a guess for the value of the hidden bit b used by the Test oracle. $\mathcal{A}$ wins the game if $b = b'$.

In the game above, the adversary can keep querying the oracles even after it asked some Test queries. However, when there was the query $\mathsf{Test}(C_i^\pi)$ asked, the adversary is prohibited from querying $\mathsf{Reveal}(C_j^\omega)$ for some $C_j^\omega \in \mathsf{pid}_i^\pi$. This restriction reflects the fact that the adversary can win the game unfairly by using the information obtained via the query $\mathsf{Reveal}(C_j^\omega)$.

Given the game above, the advantage of $\mathcal{A}$ in attacking the protocol P is defined as $\mathsf{Adv}_P(\mathcal{A}) = |2 \cdot \Pr[b = b'] - 1|$. Note that this definition is equivalent to say that the advantage of $\mathcal{A}$ is the difference between the probabilities that $\mathcal{A}$ outputs 1 in the following two experiments constituting the game: the *real experiment* where the query to the Test oracle is answered with the real session key, and the *random experiment* where the Test query is answered with a random session key. Thus, if we denote the real and the random experiments respectively as $\mathsf{Exp}_P^{\mathsf{real}}(\mathcal{A})$ and $\mathsf{Exp}_P^{\mathsf{rand}}(\mathcal{A})$, the advantage of $\mathcal{A}$ can be equivalently defined as $\mathsf{Adv}_P(\mathcal{A}) = |\Pr[\mathsf{Exp}_P^{\mathsf{real}}(\mathcal{A}) = 1] - \Pr[\mathsf{Exp}_P^{\mathsf{rand}}(\mathcal{A}) = 1]|$, where the outcomes of the experiments is the bit output by $\mathcal{A}$.

We say that the group key exchange protocol P is *secure* if $\mathsf{Adv}_P(\mathcal{A})$ is negligible for all probabilistic polynomial time adversaries $\mathcal{A}$. To quantify the security of protocol P in terms of the amount of resources expended by adversaries, we let $\mathsf{Adv}_P(t, Q)$ denote the maximum value of $\mathsf{Adv}_P(\mathcal{A})$ over all $\mathcal{A}$ with time complexity at most t and query complexity at most Q.

3 Our Protocol

We here present a password-authenticated GKE protocol n-PAKE.

Participants. The protocol participants consist of: (1) a set of clients $C_1, \ldots, C_n \in \mathcal{C}$ who wish to establish a common session key and (2) a server $S \in \mathcal{S}$ who provides the clients with a centralized authentication service. The server S is trusted to behave in an "honest but curious" manner; that is, S may attempt to learn the session key only by passive eavesdropping.

Building Blocks. The cryptographic building blocks of n-PAKE include:

- *a password-authenticated 3-party key exchange protocol* 3-PAKE that, given two users $C_i, C_j \in \mathcal{U}$ and a server $S \in \mathcal{S}$ (or rather their identities), returns either a secret key $\kappa \in \{0,1\}^k$ or a special symbol $\top$ (indicating failure of the key establishment).
- *a collision-resistant pseudorandom function family* $\mathcal{F} = \{F^\ell\}_{\ell \in \mathbb{N}}$ with $F^\ell = \{F_s^\ell\}_{s \in \{0,1\}^L}$. Collision-resistance means, informally, that there exists a value v such that no efficient adversary can find two different indices $s, s' \in \{0,1\}^L$ such that $F_s(v) = F_{s'}(v)$. We assume three publicly known values v_1, v_2 and v_3 that satisfies the collision-resistance condition.
- *a hash function* $H : \{0,1\}^* \to \{0,1\}^L$ from a family of universal hash functions. H is used to select an index within the aforementioned collision-resistant pseudorandom function family.

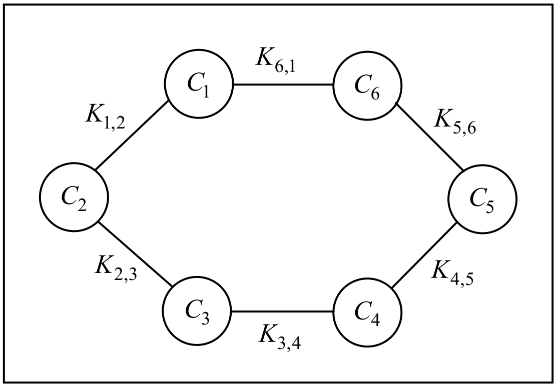


Fig. 1. Pairwise key establishments among $C_1, \ldots, C_6$

Initialization. Before n-PAKE is ever executed, the protocol participants generate public parameters and long-term secrets. The server S determines the above-mentioned building blocks and their related parameters. Each client C_i chooses a password pw_i and stores it on the server S.

Protocol Execution. If the underlying 3-party protocol 3-PAKE takes r rounds of communications, then the protocol n-PAKE takes $r + 2$ rounds. Recall that each C_i receives as input a group ID $\mathsf{gid}_i = \{C_1, \ldots, C_n\} \cup \{S\}$ to start to run the protocol. n-PAKE runs as follows (throughout the protocol description, all indices are to be taken in a cycle, i.e., $C_{n+1} = C_1$, etc.):

[ROUND $1 \sim r$]: Each neighboring pair of clients C_i and C_{i+1}, for $i = 1, \ldots, n$, generates a pairwise key $K_{i,i+1}$ by running 3-PAKE with the server S. Accordingly, at the end of the r^{th} round, each C_i holds two pairwise keys $K_{i-1,i}$ and $K_{i,i+1}$ shared respectively with C_{i-1} and C_{i+1}. (Fig. 1 shows pairwise key establishments among $C_1, \ldots, C_6$.)

[ROUND $r + 1$]: Each C_i computes

$$\overleftarrow{\sigma}_i = F_{H(C_i \| K_{i-1,i} \| \mathsf{gid}_i)}(v_1),$$
$$\overrightarrow{\sigma}_i = F_{H(C_i \| K_{i,i+1} \| \mathsf{gid}_i)}(v_1),$$

and sends $\overleftarrow{\text{AUTH}}_i = \langle C_i \| \overleftarrow{\sigma}_i \rangle$ and $\overrightarrow{\text{AUTH}}_i = \langle C_i \| \overrightarrow{\sigma}_i \rangle$ respectively to C_{i-1} and C_{i+1}. Upon receiving $\overrightarrow{\text{AUTH}}_{i-1}$ and $\overleftarrow{\text{AUTH}}_{i+1}$, C_i verifies the correctness of both $\overrightarrow{\sigma}_{i-1}$ and $\overleftarrow{\sigma}_{i+1}$ in the straightforward way. If either one of the verifications fails, C_i aborts the protocol.

[ROUND $r + 2$]: Each C_i computes

$$X_i = K_{i-1,i} \oplus K_{i,i+1}.$$

and broadcasts $\text{XOR}_i = \langle C_i \| X_i \rangle$. Upon receiving the XOR-values, C_i checks that $X_1 \oplus X_2 \oplus \cdots \oplus X_n = 0$. If the check fails, C_i aborts the protocol.

[Key Computation]: Using $K_{i-1,i}$ and the XOR-values, each C_i computes

$$\begin{aligned} K_{i-2,i-1} &= X_{i-1} \oplus K_{i-1,i}, \\ K_{i-3,i-2} &= X_{i-2} \oplus K_{i-2,i-1}, \\ &\vdots \\ K_{i,i+1} &= X_{i+1} \oplus K_{i+1,i+2}. \end{aligned}$$

Then, C_i defines a master key

$$K = \langle K_{n,1} \| K_{1,2} \| \ldots \| K_{n-1,n} \| \mathsf{gid}_i \rangle,$$

computes the session key $SK_i = F_{H(K)}(v_2)$, and sets $\mathsf{sid}_i = F_{H(K)}(v_3)$.

For the instantiation of 3-PAKE, any particular choice that is, informally speaking, secure against an active adversary will do. Possible candidates include the generic 3-party protocols from [3] and [26].

Remark 1. We assume that the number of clients participating in n-PAKE is greater than 2, i.e., $n \geq 3$. It is because in case of $n = 2$, running the 3-party protocol 3-PAKE suffices for two clients to establish a session key.

Efficiency Consideration. Our 3-to-n compiler is quite efficient both in terms of computation cost and communication cost. The transformation process of the compiler adds $O(n)$ computation complexity and $O(1)$ round complexity to those of the underlying 3-party protocol. From the practical point of view, the $O(n)$ increase in computation overhead is not that significant because the increase is purely due to exclusive-or (XOR) operations, which can be implemented efficiently in hardware and/or software.

Security Result. Here we claim that the GKE protocol n-PAKE is secure against active adversaries under the security of the 3-party key exchange protocol 3-PAKE against active adversaries. The following theorem makes this claim precise.

Theorem 1. *Let $Q = (q_{\text{exec}}, q_{\text{send}}, q_{\text{reve}}, q_{\text{corr}}, q_{\text{test}})$ and $Q' = (n \cdot q_{\text{exec}}, 2 \cdot q_{\text{send}}, q_{\text{reve}}, q_{\text{corr}}, q_{\text{test}})$. For any adversary with time complexity at most t and query complexity at most Q, its advantage in breaking the security of protocol n-PAKE is upper bounded by:*

$$\mathsf{Adv}_{n\text{-}PAKE}(t, Q) \leq \mathsf{Adv}_{3\text{-}PAKE}(t', Q'),$$

where $t' = t + O(n \cdot q_{\text{exec}} t_{3\text{-}PAKE} + 2 \cdot q_{\text{send}} t_{3\text{-}PAKE})$ and $t_{3\text{-}PAKE}$ is the time required for execution of 3-PAKE by any party.

References

1. Abdalla, M., Bohli, J.-M., González Vasco, M.I., Steinwandt, R. (Password) authenticated key establishment: From 2-party to group. In: Vadhan, S.P. (ed.) TCC 2007. LNCS, vol. 4392, pp. 499–514. Springer, Heidelberg (2007)

2. Abdalla, M., Bresson, E., Chevassut, O., Pointcheval, D.: Password-based group key exchange in a constant number of rounds. In: Yung, M., Dodis, Y., Kiayias, A., Malkin, T. (eds.) PKC 2006. LNCS, vol. 3958, pp. 427–442. Springer, Heidelberg (2006)
3. Abdalla, M., Fouque, P.-A., Pointcheval, D.: Password-based authenticated key exchange in the three-party setting. In: Vaudenay, S. (ed.) PKC 2005. LNCS, vol. 3386, pp. 65–84. Springer, Heidelberg (2005)
4. Boyd, C., Nieto, J.: Round-optimal contributory conference key agreement. In: Desmedt, Y.G. (ed.) PKC 2003. LNCS, vol. 2567, pp. 161–174. Springer, Heidelberg (2002)
5. Bresson, E., Chevassut, O., Pointcheval, D.: Group diffie-hellman key exchange secure against dictionary attacks. In: Zheng, Y. (ed.) ASIACRYPT 2002. LNCS, vol. 2501, pp. 497–514. Springer, Heidelberg (2002)
6. Bresson, E., Chevassut, O., Pointcheval, D., Quisquater, J.-J.: Provably authenticated group Diffie-Hellman key exchange. In: 8th ACM Conference on Computer and Communications Security (CCS 2001), pp. 255–264 (2001)
7. Burmester, M., Desmedt, Y.G.: A secure and efficient conference key distribution system. In: De Santis, A. (ed.) EUROCRYPT 1994. LNCS, vol. 950, pp. 275–286. Springer, Heidelberg (1995)
8. Burmester, M., Desmedt, Y.: Efficient and secure conference-key distribution. In: Lomas, M. (ed.) Security Protocols 1996. LNCS, vol. 1189, pp. 119–129. Springer, Heidelberg (1997)
9. Byun, J.W., Lee, D.-H.: N-party encrypted diffie-hellman key exchange using different passwords. In: Ioannidis, J., Keromytis, A.D., Yung, M. (eds.) ACNS 2005. LNCS, vol. 3531, pp. 75–90. Springer, Heidelberg (2005)
10. Byun, J.W., Lee, S.-M., Lee, D.-H., Hong, D.: Constant-round password-based group key generation for multi-layer ad-hoc networks. In: Clark, J.A., Paige, R.F., Polack, F.A.C., Brooke, P.J. (eds.) SPC 2006. LNCS, vol. 3934, pp. 3–17. Springer, Heidelberg (2006)
11. Choo, K.-K.R., Boyd, C., Hitchcock, Y.: Errors in computational complexity proofs for protocols. In: Roy, B. (ed.) ASIACRYPT 2005. LNCS, vol. 3788, pp. 624–643. Springer, Heidelberg (2005)
12. Facebook, http://www.facebook.com
13. Ingemarsson, I., Tang, D., Wong, C.: A conference key distribution system. IEEE Transactions on Information Theory 28(5), 714–720 (1982)
14. Katz, J., Shin, J.: Modeling insider attacks on group key-exchange protocols. In: 12th ACM Conference on Computer and Communications Security (CCS 2005), pp. 180–189 (2005)
15. Katz, J., Yung, M.: Scalable protocols for authenticated group key exchange. In: Boneh, D. (ed.) CRYPTO 2003. LNCS, vol. 2729, pp. 110–125. Springer, Heidelberg (2003)
16. Kim, H., Lee, S., Lee, D.: Constant-round authenticated group key exchange for dynamic groups. In: Lee, P.J. (ed.) ASIACRYPT 2004. LNCS, vol. 3329, pp. 245–259. Springer, Heidelberg (2004)
17. Kim, Y., Perrig, A., Tsudik, G.: Simple and fault-tolerant key agreement for dynamic collaborative groups. In: 7th ACM Conference on Computer and Communications Security (CCS 2000), pp. 235–244 (2000)
18. Kim, Y., Perrig, A., Tsudik, G.: Communication-efficient group key agreement. In: IFIP SEC 2001, pp. 229–244 (2001)

19. Kwon, J., Jeong, I., Sakurai, K., Lee, D.: Password-authenticated multi-party key exchange with different passwords. Cryptology ePrint Archive, Report 2006/476 (2006)
20. Mayer, M., Yung, M.: Secure protocol transformation via "Expansion": From two-party to groups. In: 6th ACM Conference on Computer and Communications Security (CCS 1999), pp. 83–92 (1999)
21. Nam, J., Paik, J., Kim, U.-M., Won, D.H.: Constant-round authenticated group key exchange with logarithmic computation complexity. In: Katz, J., Yung, M. (eds.) ACNS 2007. LNCS, vol. 4521, pp. 158–176. Springer, Heidelberg (2007)
22. Nam, J., Paik, J., Kim, U., Won, D.: Security enhancement to a password-authenticated group key exchange protocol for mobile ad-hoc networks. IEEE Communications Letters 12(2), 127–129 (2008)
23. Pereira, O., Quisquater, J.-J.: A security analysis of the Cliques protocols suites. In: 14th IEEE Computer Security Foundations Workshop, pp. 73–81 (2001)
24. Shim, K., Woo, S.: Cryptanalysis of tripartite and multi-party authenticated key agreement protocols. Information Sciences 177(4), 1143–1151 (2007)
25. Twitter, http://twitter.com
26. Wang, W., Hu, L.: Efficient and provably secure generic construction of three-party password-based authenticated key exchange protocols. In: Barua, R., Lange, T. (eds.) INDOCRYPT 2006. LNCS, vol. 4329, pp. 118–132. Springer, Heidelberg (2006)
27. Yi, X., Tso, R., Okamoto, E.: ID-Based group password-authenticated key exchange. In: Takagi, T., Mambo, M. (eds.) IWSEC 2009. LNCS, vol. 5824, pp. 192–211. Springer, Heidelberg (2009)

Does Privacy Information Influence Users' Online Purchasing Behavior?

Jimmy H. Nguyen and Kim-Phuong L. Vu

California State University Long Beach 1250 Bellflower Blvd., Long Beach, CA
Jimmy.Nguyen@student.csulb.com, kvu8@csulb.edu

Abstract. Web sites provide privacy policies to inform users about how their personal information is being handled. However, privacy policies are usually difficult to find or are written in a legal language that is hard for the average user to understand. Thus, privacy information is often ignored by users. The goal of the present study was to determine whether more salient presentation of privacy information, through a summary provided by the Privacy Finder Web site, would influence user purchasing behavior. Specifically, we examined whether Privacy Finder influenced college students' comfort levels in making small and large purchases from familiar and unfamiliar Web sites in a simulated e-commerce task. Users were more comfortable purchasing inexpensive items, as well as making purchases from more familiar Web sites. However, Privacy Finder did not influence their purchasing behaviors or comfort levels for the different types of Web sites.

Keywords: Web Privacy, Privacy Finder, E-Commerce, Usability.

1 Introduction

Many Internet users that surf the Web daily are unaware of the potential risks associated with their normal Web activities such as making online purchases, visiting social networking sites, booking hotels/flights, or even transferring files. What is at risk is the user's privacy. A user's personal information may be constantly changing hands or, in some cases servers, without notice or permission. A user's information can even be sold to other parties for profit or market research. Proctor, Ali, and Vu [1] showed that most Web sites providing a service require that a user provide his/her name (97%), e-mail address (71%), and physical address (88%). Additionally, sites that Internet users visit can collect information about their behaviors such as type of purchase, Web sites or pages visited, and other information that can result in the profiling of the user. This may strike Internet users as surprising, but there is no federal law protecting users from having their personal information shared. Federally, there are specific privacy laws in place such as: (a) the Gramm-Leach-Bliley Act, which protects a user's financial information that is stored by financial institutions; (b) HIPAA, which protects individually identifiable health information; (c) the Children's Online Privacy Protection Act, which allows parents to decide what information can be collected on their children and how that information is used.

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 349–358, 2011.

However, these laws do not apply to most users conducting everyday transactions on the Web. To combat consumer privacy invasion issues and reassure users regarding privacy concerns, many Web sites have posted privacy policies that state how user information is reportedly handled by the host organization. Even though privacy policies are provided, Internet users, including some that express major privacy concerns, do not often view organizations' privacy policies [2]. This lack of concern is alarming given that users should know how their personal information is being used.

One reason why users do not read privacy policies is that they have a common misconception about privacy policies [2, 3]. According to Turow et al. [3], a majority of American users feel that they are protected from their data being sold or from fraud based on the mere presence of a privacy policy link. Because of these cues, Internet users usually "accept" terms of a site without reading or even briefly reviewing them. Moreover, users tend to trust Web sites that they frequently visit, have previously done business with, or are associated with popular brand names [4]. Users have also been shown to ignore privacy policies because they feel that a Web site is trustworthy or because they are confusing, or excessively long [5]. Even if users have the intention to read a privacy policy, searching for the policy itself can be an inconvenient task. Byers, Cranor, and Kormann [6] showed that the time it takes to locate a site's privacy policy may even force users to forgo the search altogether. Neglecting a Web site's privacy policy, though, may lead to undesirable consequences. For example, not having an understanding of what a Web site does with the user's data can lead to annoyances such as spam or junk mail. Web sites can also monitor the user's behaviors such as frequently searched topics or patterns of usage, and use those data to identify the user for research or marketing purposes.

1.1 Behavior: Stated vs. Observed

Users also report that they base decisions regarding whether to use a Web site based on its privacy policy [7]. Yet, when observed and tested, users' reported behaviors and observed behaviors are discrepant. Findings from several studies show that most users rarely visit privacy policy pages and even fewer read the policies [2, 9]. Users also report that privacy issues are the main reason for not purchasing from the Internet [7, 8], but the cost of an item can nullify concern for privacy. For example, Vu et al. [5] showed that users would make a purchase on an unfamiliar site if the price was low. Users readily provided credit card information to make small purchases, not realizing that the host site would obtain their credit card information regardless of the size of the purchase.

1.2 Comprehension and Readability of Privacy Policies

The way a Web site handles a user's information is stated in its privacy policy, which describes how the information being collected is stored and used by the site and its host organization. This allows users to view a site's information practices as well as other pertinent facts. However, policies are often written using legal terms and language, so if users review a site's privacy terms, they may not understand it. Jenson and Potts [9] analyzed the readability of 47 high traffic sites' privacy policies, ranging from financial sites to retail sites, and found that most policies required an education grade level of 12 or higher. In another study, Proctor et al. [8] analyzed 100 Web sites' privacy policies using readability tests such as the Flesch Grade Level (FGL), which provides a grade level required to comprehend a text. For Web sites such as banks, insurance companies, retail stores, and pharmacies, the FGL was at least 13.

The comprehensibility of a Web site's privacy policy is important, but studies have shown that even among users with a college level education, many are not able to understand certain aspects of a Web site's policy [11]. Comprehension is even more critical when sites that are used require personal information from users such as medical records, credit card numbers, and bank accounts. Not being able to understand a privacy policy can put one's personal information at risk. Finally, the length and organization of privacy policies can also influence whether users read a policy. Users tend to avoid reading long policies. Moreover, long policies can even mislead users into believing that there are more privacy assurances contained in them than those provided by shorter policies [8].

1.3 Checking Privacy Preferences

A proposed solution for increasing the usefulness of privacy policies is to have guidelines and standards regarding their layout and content. The development of a privacy policy standard, called the Platform for Privacy Preferences (P3P), has been implemented in some Web sites since 2000. P3P's purpose is to make privacy policies more accessible to all Internet users by making the format universal and easier to navigate and search. It has also allowed software to be developed that can "read" a privacy policy for users. One software program designed to read privacy policies is called Privacy Bird®. Privacy Bird allows users to select or customize a privacy setting, scan a Web site's privacy policy, and review the results. Cranor et al.'s [12] research on Privacy Bird has shown that it could increase users' awareness of privacy policies by increasing the number of users reading a privacy policy after visiting a Web site that received a violation alert. Users also reported that they provided less information to sites that violated their privacy preferences, stopped visiting those sites, and tried to visit sites with better policies as a result of using Privacy Bird [12].

Although Privacy Bird attempted to provide a bridge between users and usable privacy policies, there were some problems with the program and interface that hindered its adoption and use. Proctor et al. [1] showed that participants incorrectly set Privacy Bird's configurations when intending to check for specific privacy concerns. In addition, users thought that they had set Privacy Bird to perform protection tasks that it could not do. Another problem is the lack of widespread implementation of P3P on the Internet. Many users felt that Privacy Bird would be more useful if P3P became more widely adopted [13]. Finally, Privacy Bird requires users to download and install the program, which limits its use on public computers.

An alternative to Privacy Bird is Privacy Finder, which does not require any downloading or installing. Privacy Finder (www.privacyfinder.org) is a search engine with privacy filters. Working in similar fashion to Privacy Bird, users are able to go to the Privacy Finder site and select their privacy settings to be low, medium, high, or custom. Users can then input the site that they wish to enter, and the results of the search appear along with a link to the page. Privacy Finder indicates the level of compliance with 4 boxes next to the Web site link (see Figure 1). If a site does not meet the selected settings, a link to reasons why it does not is also available for the user to review. This enables the user to decide whether to proceed. Although Privacy Finder requires users to visit an additional site to determine whether the intended site is compatible with their privacy preferences, a downloadable version is available to install into Web browsers such as Firefox and Internet Explorer. Initial research conducted on Privacy Finder by its developers has shown that, when privacy policy information is made more salient to the users, users are more likely to purchase from Web sites that have better protection and will even pay more money to purchase from Web sites with high privacy policy compatibility [14].

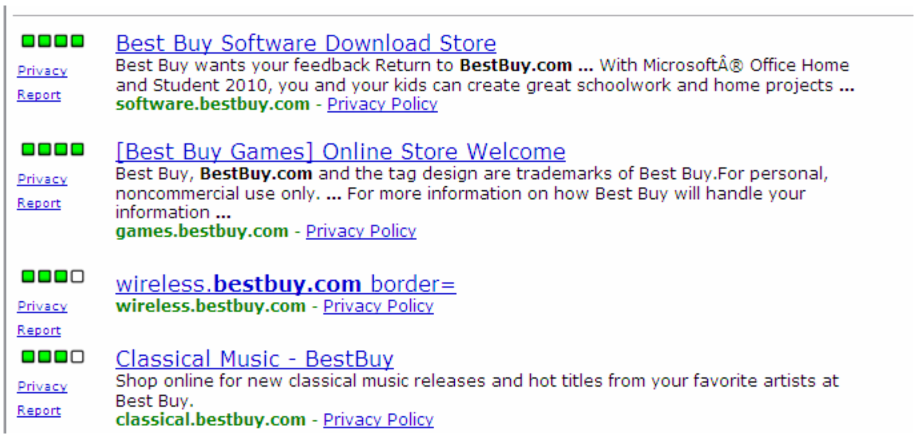


Fig. 1. Screen shot illustrating Privacy Finders' results for a search of bestbuy.com

1.4 Present Study

The present study further examined the effect of Privacy Finder's search engine results on users' purchasing behavior. Participants were asked to complete a simulated e-commerce task using Privacy Finder. This study also compared users' behavior when using sites with which they were familiar to sites with which they were not. We hypothesized that users would be more likely to purchase items from familiar sites, sites with which users had more experience, and sites that were more reputable. Also, we hypothesized that familiarity with a site would affect how users configured privacy settings on Privacy Finder and that Privacy Finder's warnings would influence users' purchasing decisions as well as comfort levels. In addition, we compared purchases of expensive versus inexpensive products, because price is a factor known to influence users' buying decisions [5].

2 Method

2.1 Participants

A total of 16 participants completed the experiment. Participants ranged from 18 to 27 years of age, with mean age of approximately 21 years. Participants' education level ranged from college freshman to graduate student, with most being sophomores.

The participants reported that they used the Internet very often (M = 1.06, with 1 being very often and 4 being never) and were very experienced with computers (M = 1.6, with 1 being very experienced and 4 being no experience). However, participants reported not being very knowledgeable about privacy policies (M = 2.3, with 1 being very knowledgeable and 4 being not knowledgeable) and not having much experience reading them (M = 3, with 1 being very experienced and 4 being no experience).

2.2 Materials and Apparatus

Participants used Privacy Finder, with Internet Explorer 8, to carry out simulated e-commerce tasks on a personal computer running Windows XP Professional (Pentium 4, 3.20 GHz). Camtasia Studio Version 2 was used to record all of the participants' computer actions on a 17" LCD monitor.

Participants were each given two scenarios, one at a time, with specific tasks to be completed. The scenarios stated, "Your grandmother is interested in making several purchases online. However, she is unfamiliar with the Internet and trusts you with her information to make these purchases. She would like you to help her with the following purchases."

Scenario 1: Your grandmother wants to purchase the book "Dear John" by Nicholas Sparks for your cousin's 17th birthday. She plans on spending between $10 and $20 for his present. You will be given Web sites one at a time in order to decide whether or not you would make the purchase at each given site. Feel free to look at any of the information that is available to you on each site in order to assist you in making your decision. Once you've determined whether you would make the purchase from a site, you may stop and I will ask you a few short questions.

Scenario 2: Your grandmother wants to purchase a 40" – 50" HDTV for your parents as a Christmas gift. She can only afford to spend under $1000 for it. You will be given Web sites one at a time in order to decide whether or not you would make the purchase at each given site. Feel free to look at any of the information that is available to you on each site in order to assist you in making your decision. Once you've determined whether you would make the purchase from a site, you may stop and I will ask you a few short questions.

A fake profile containing personal and financial information was also provided for the participant to use when completing their task to increase the laboratory tasks' realism. In addition, a list of pre-determined e-commerce sites was given to participants, who visited each site to determine whether a purchase was to be made using that particular site. The following sites were presented to the participant for the book purchasing task: www.barnesandnobles.com, www.borderstores.com, www.amazon.com, www.alibris.com, www.abebooks.com. For the HDTV purchasing task, the sites used were: www.bestbuy.com, www.frys.com, www.amazon.com, www.jr.com, www.savinglots.com. For each task, the order of the sites was varied between participants.

2.3 Procedure

Consent forms were given to participants to complete, and the experimenter briefly explained the purpose of the study verbally. Once participants signed the consent forms, they were given a demographic questionnaire to complete and return. Participants were then reminded that the experiment would last approximately 2 hours and were told that they could stop their participation at any time without penalty.

Participants were then led into an adjacent lab room where they were seated in front of a personal computer. Here, participants were given a description of Privacy Finder and a demonstration of how it worked. Participants were told that Privacy Finder allowed them to check a Web site's privacy policy while searching normally. Participants were also allowed to test Privacy Finder by setting their own privacy preferences and searching for a Web site of their choice. Then, they were instructed to complete the tasks and encouraged to think aloud.

The first task was to search for a book and decide whether they would purchase it from each of several e-commerce sites with which they were provided. The second task involved the purchase of an HDTV from several e-commerce sites that were also provided to the participants. They were allowed to adjust their privacy settings once they were given the name of each site that was to be visited. After each task was completed, participants were asked several follow-up questions regarding how much experience they had with each e-commerce site, how they felt about each site's security, and how comfortable they felt in making a purchase at each site. After all follow up questions were answered, a questionnaire about general Internet behaviors was administered and completed.

3 Results

3.1 Familiarity with Web Sites

Familiarity with a site was determined by examining participants' reported experience with the different Web sites. Experience was determined by a 7-point Likert scale with, 1 being "No Previous Experience" and 7 being "Extensive Previous Experience". An analysis of variance (ANOVA) on experience ratings as a function of e-commerce Web site was carried out. For the book task, there was a significant effect of Web site, $F(1, 15) = 11.82$, $p < .001$, see Figure 2. Participants were more familiar with Amazon, Borders, and Barnes and Noble than with Abe Books and Alibris. Results for the HDTV task also showed a significant effect of Web site, $F(1, 15) = 21.29$, $p < .001$. Participants reported being more familiar with Best Buy and Amazon than with Fry's, JR, and Saving Lots. Based on these findings, the Web sites were grouped into 2 categories of familiar and unfamiliar.

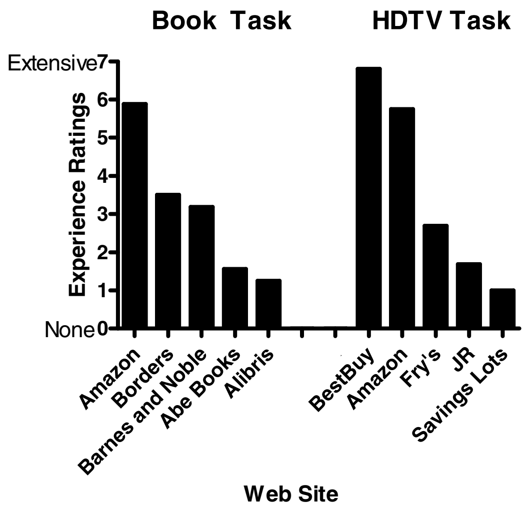


Fig. 2. Familiarity ratings for each Web site for the Book and HDTV tasks

A chi-square test was then conducted to examine whether familiarity influenced the sites from which users would make their purchases for the two tasks. Participants indicated that they would make more purchases from familiar than unfamiliar sites for both the book task, $\chi 2$ (N=16) = 6.25, $p < .02$, and TV task, $\chi 2$ (N=16) = 12.25, $p < .001$.

3.2 Comfort Ratings

A 2 Familiarity (Familiar or Unfamiliar) X 2 Cost [low (Book task) or high (HDTV task)] repeated-measures ANOVA was conducted on mean comfort ratings for purchasing an item. The main effect of cost was significant, $F(1, 15) = 7.33$, $p < .05$. Participants reported that they were more comfortable making inexpensive purchases ($M = 5.35$) than expensive purchases ($M = 4.39$). The main effect of familiarity, $F(1, 15) = 74.92$, $p < .001$, indicates that participants were more comfortable in making purchases at familiar sites ($M = 6.84$) than unfamiliar sites ($M = 2.93$). The interaction between familiarity and cost was also significant, $F(1, 15) = 16.57$, $p < .001$. This interaction reflected the pattern that although participants were more comfortable making purchases at familiar sites than unfamiliar sites, the difference in comfort ratings was much greater when the purchase was expensive ($M = 7.00$ vs. 1.79) than when it was inexpensive ($M = 6.69$ vs. 4.10).

3.3 Security Ratings

Participants were asked to rate how they felt about a site's security level, after making a purchase of an inexpensive or expensive item, on a 10-point Likert scale, with 1 being not secure at all and 10 being very secure. A paired samples t test was conducted but no significant effect was obtained, indicating that users did not report any difference in how they felt about security of the site when purchasing either item, $t(15) = -1.52$, $p > .15$. In general, users thought that both the TV sites ($M = 8.13$) and the book sites ($M = 7.63$) were secure.

3.4 Influence of Privacy Finder

A chi-square test was conducted to examine whether more participants set their privacy settings to high, medium, low or custom, based on whether a site was familiar or unfamiliar. No effects were significant. Most users set their privacy settings to the medium level regardless of the purchasing task. Because the results provided by Privacy Finder depended on the user setting, we tabulated the frequency (see Table 1) with which participants indicated that they would make a purchase from a familiar versus unfamiliar site as a function of Privacy Finder's warnings (no violation, unknown, violation), but did not conduct any formal analyses.

Table 1. Number of purchases as a function of Privacy Finder warnings for each task

	Privacy Finder Warning		
Familiarity	No Violation	Unknown	Violation
		Book Purchase	
Familiar	5	3	5
Unfamiliar	0	3	0
		HDTV Purchase	
Familiar	0	11	3
Unfamiliar	0	2	0

Privacy Finder did not appear to influence purchasing behaviors beyond the effects of familiarity. In fact, participants' comfort ratings for making the book purchase from familiar sites were similar regardless of Privacy Finder's warnings (M = 6.4 for no violation, M = 6.4 for unknown, and M = 7.3 for violation), and these ratings were higher than that for unfamiliar sites (M = 4.1).

4 Discussion

The goal of the present study was to examine whether users' purchasing behavior changes for familiar versus unfamiliar sites when privacy policies are brought into focus with the Privacy Finder search tool. We found that familiarity with a site is an important factor influencing whether users would make a purchase. For both the book and HDTV task, more users indicated that they would make the purchase from a familiar rather than unfamiliar Web site. In addition, users reported feeling more comfortable when making a purchase from a familiar than unfamiliar site. The size of the purchase also influenced user behaviors in that users were more comfortable making a smaller (book) purchase than a larger (HDTV) purchase. These findings are consistent with past research [11].

The type of site or familiarity with a site did not influence how users set their privacy setting on Privacy Finder. Users tended to set the privacy setting at medium regardless of type of site or regardless of their familiarity with the site. It should be noted that the medium setting was the default setting, and that users seldom changed the settings after doing so initially. In other words, users did not set stricter criteria for sites from which they would make larger purchases or for unfamiliar sites. More importantly, Privacy Finder did not seem to influence user's purchasing behaviors. For the book task, users were as likely to make purchases from a familiar site when its privacy policy violated their privacy preferences as when it did not. One reason for the lack of an influence of Privacy Finder is that users indicated that they felt that all Web sites were secure and that their transactions were safe. As such, participants continued their purchasing task without paying much attention to the results of Privacy Finder. This pattern of behavior matches those of other studies and is captured by Berendt et al.'s [7] statement, "Once they are in an online interaction, they often do not monitor and control their actions sufficiently; privacy statements seem to have no impact on behavior" (p. 105).

4.1 Limitations

We close by noting limitations of the present study. First, the lack of influence of Privacy Finder may be limited to the sample used in the study. The sample was drawn from college students at a public university. Although college students are e-commerce users and are thus part of the targeted population, they may be less concerned about security and privacy than other types of users. Thus, these findings need to be replicated with a larger sample size of more diverse user groups. However, the finding that college students may not be concerned about privacy is one that should be noted, as many companies target college students. Second, participants may have been less concerned about security and privacy than they would be otherwise because the study used a simulated e-commerce task in which the users were not spending their own money or providing their personal information. Users may be more concerned if they are actually providing their own credit card information. Finally, we did not control for the privacy level setting in the present study. Rather, participants were allowed to set their privacy level to match their preference. As such, we could not systematically compare the Privacy Finder warnings across the different sites.

References

1. Proctor, R.W., Vu, K.-P.L., Ali, M.A.: Usability of User Agents for Privacy-Preference Specification. In: Smith, M.J., Salvendy, G. (eds.) HCII 2007. LNCS, vol. 4558, pp. 766–776. Springer, Heidelberg (2007)
2. Jensen, C., Potts, C., Jensen, C.: Privacy practices of Internet users: Self-reports versus observed behavior. Int. J. Hum.-Comput. St. 63, 203–227 (2005)
3. Turow, J., Feldman, L., Meltzer, K.: Open to exploration: American shopper online and offline. A report from the Annenberg public policy center of the University of Pennsylvania. University of Pennsylvania (2005)
4. Metzger, M.J.: Effects of site, vendor, and consumer characteristics on Web site trust and disclosure. Commun. Res. 33, 155–179 (2006)
5. Vu, K.-P.L., Chambers, V., Garcia, F.P., Creekmur, B., Sulaitis, J., Nelson, D., Pierce, R., Proctor, R.W.: How Users Read and Comprehend Privacy Policies. In: Smith, M.J., Salvendy, G. (eds.) HCII 2007. LNCS, vol. 4558, pp. 802–811. Springer, Heidelberg (2007)

6. Byers, S., Cranor, L.F., Kormann, D.: Automated analysis of P3P-enabbled Web sites. In: ICES 2003, Pittsburgh, Pennsylvania, pp. 326–338 (2003)
7. Berendt, B., Gunther, O., Spiekermann, S.: Privacy in e-commerce: Stated preference vs. actual Behavior. Commun. ACM 48, 101–106 (2005)
8. Jupiter Research, `http://www.ftc.gov/bcp/workshops/security/020520leathern.pdf`
9. Jensen, C., Potts, C.: Privacy policies as decision-making tools: An evaluation of online privacy notices. In: CHI 2004, New York, NY, vol. 6, pp. 471–478 (2004)
10. Proctor, R.W., Ali, M.A., Vu, K.-P.L.: Examining usability of Web privacy policies. Int. J. Hum.-Comput. Int. 24, 307–328 (2008)
11. Vu, K.-P.L., Garcia, F.P., Nelson, D., Sulaitis, J., Creekmur, B., Chambers, V., Proctor, R.W.: Examining User Privacy Practices While Shopping Online: What Are Users Looking for? In: Smith, M.J., Salvendy, G. (eds.) HCII 2007. LNCS, vol. 4558, pp. 792–801. Springer, Heidelberg (2007)
12. Cranor, L.F., Arjula, M., Guduru, P.: Use of a P3P user agent by early adopters. In: WPES 2002, ACM Workshop on Privacy in the Electronic Society, New York, NY, pp. 1–10 (2002)
13. Cranor, L.F., Guduru, P., Arjula, M.: User interfaces for privacy agents. ACM Trans. Comput.-Hum. Interact. 13, 135–178 (2006)
14. Tsai, J., Egelman, S., Cranor, L.F., Acquisti, A.: The effect of online privacy information on purchasing behavior: An experimental study. In: The 6th Workshop on the Economics of Information Security, WEIS (2008), `http://weis2007.econinfosec.org/papers/57.pdf`

Analysis of Authentication Protocols with Scyter: Case Study

Očenášek Pavel

Faculty of Information Technology,
Brno University of Technology,
Božetěchova 2, 612 66 Brno, Czech Republic
ocenaspa@fit.vutbr.cz
http://www.fit.vutbr.cz/~ocenaspa/

Abstract. This paper describes one of possible implementations of method used for protocol design described in a paper from 1998 by L. Buttyán, S. Staamann and U. Wilhelm which proposes a backward search (regression) when synthesizing an authentication protocol. Furthermore, the approach presented in this paper allows a designer to define participants who are trustworthy enough to transfer information between other two participants without existence of a direct channel to achieve basic routing functionality.

Keywords: Logic, Authentication, Key Distribution, Security Protocol, Routing.

1 Introduction

Communication between the various types of electronic devices from computers to mobile phones to the payment terminal is one of the basic requirements of today. Transfer information between devices is needed, however, clearly defined rules. These rules are collectively referred to as a communication protocol. Information transmitted using these protocols through the network may have high value and are now under attack. Therefore represent a special area of security protocols deployed at places where requires a secure, encrypted communications. These are the protocols carrying sensitive information such as payment orders or authentication data. Design, testing and proper understanding of communication and security protocols to requires specific tools and environment to enable their simulation. Some instruments can detect possible errors or security vulnerabilities in the draft.

This paper presents the partial results of the research published within the thesis [4].

Security protocols are a separate area of communication protocols. They are mainly used in situations where we need to transmit sensitive information such as secret keys, authentication data or transactions on your bank account. These protocols provide the mechanisms using cryptographic (encryption), the basic security features

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 359–365, 2011.

such as authentication, authorization, integrity, confidentiality and sometimes unquestionable. Security features can be applied to each of the layers of the reference ISO / OSI model.

2 Scyter

Scyther [1] is a free downloadable tool designed primarily for verification and analysis of security protocols. It is based on an algorithm that provides a condensed representation (infinite) set of traces. This tool helps in analyzing the classes of possible attacks and protocol behavior. Tool to prove the correctness of indeterminate number of connections. It can operate in parallel with several (sub) protocols.

The program is written in Python and has a control console and graphical interface. The expected entry is the protocol description in the SPDL language. The program can determine whether they complied with safety requirements in the protocol if the safety requirements to generate and subsequently verified. It can analyze the protocol interaction by participants.

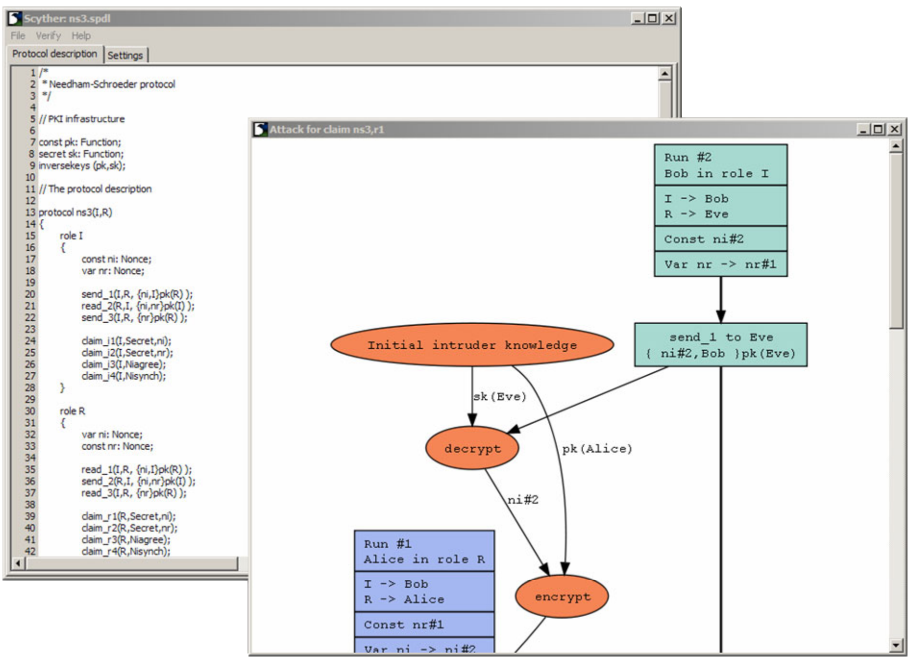


Fig. 1. Scyther GUI with the protocol example (Needham Schroeder protocol)

The program can work interactively and can not be changed at run-time conditions. It is necessary to modify the source file (eg, the built-in editor) and then run the analysis. Due to its speed, clearness and simplicity fits as a teaching aid.

3 Needham-Schroeder Public Key Protocol

As a case study, we use a version of the Needham-Schroeder protocol [2] based on asymmetric cryptography. It provides mutual entity authentication using a trusted key server and public keys. The server distributes keys upon request. The protocol was designed to communications over insecure network. The authors are Roger Needham and Michael Schroeder.

1. $A \rightarrow S: A, B$
2. $S \rightarrow A: \{ K_B, B \}_{K_S^{-1}}$
3. $A \rightarrow B: \{ N_A, A \}_{K_B}$
4. $B \rightarrow S: B, A$
5. $S \rightarrow B: \{ K_A, A \}_{K_S^{-1}}$
6. $B \rightarrow A: \{ N_A, N_B \}_{K_A}$
7. $A \rightarrow B: \{ N_B \}_{K_B}$

Fig. 2. Needham-Schroeder Public Key Protocol: Common syntax

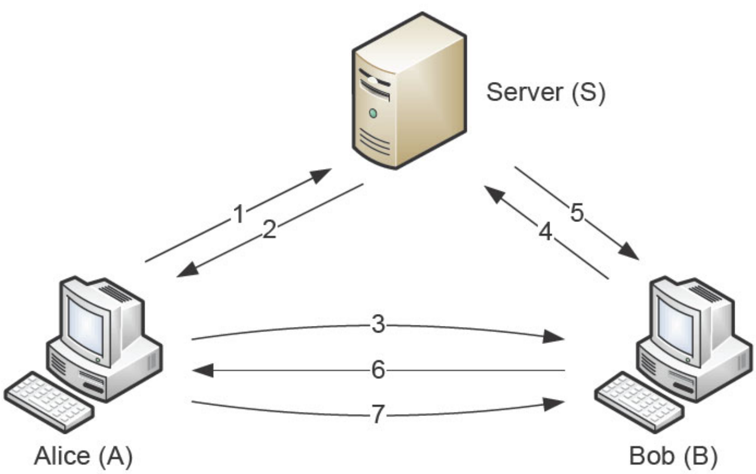


Fig. 3. Needham-Schroeder: graphical representation

4 Implementation in Scyther

The tool Scyther uses its own language SPDL (Security Protocol Language Descritpion).

We start with the keys declaration:

```
const pk: Function,
secret sk: Function,
inversekeys(pk,sk); // Pair of asymmetric keys
```

where const means the global constant. The specification of protocol behavior follows:

```
protocol NeedhamSchroederPK(I,R,S) {}
```

Within the Needham Schroeder protocol, we have to define roles for all the participating subjects (I, R, S). For example, we include the definition of role I:

```
role I {
const Ni: Nonce;    // Global variable
var Nr: Nonce;     // Local variable
send_1(I,S,(I,R)); // Message from I to S
                          // includes (I,R)
read_2(S,I, {pk(R), R}sk(S));
...
claim_I1(I,Secret,Ni);
claim_I2(I,Secret,Nr);
claim_I3(I,Nisynch);
}
```

In the final part of the protocol specification, we include the trustworthy participants of the communication, those participant we cannot trust and secret data.

```
const Alice,Bob,Server,Compromised: Agent;
untrusted Compromised;
const nc: Nonce;
compromised sk(Compromised);
```

The SPDL language is case sensitive. The comments can be written in different way (like in C language). One-line comment # or //, multi-line comments starts with /* and end with */ . The whitechars (newline, space) are ignored.

5 Verification Results

The protocol specification in SPDL language is an input for Scyther verification.

If we have a final protocol described and the main parameters are set, we can start the verification in the menu: *Verify / Verify protocol (F1)*. A window appears, where each row represents one requirement (or claim). On the line the information is listed in the order: name of protocol, role (party identification), ID requirements, and the request type parameter, informing the state to find an attack, and the button to display graphic attack.

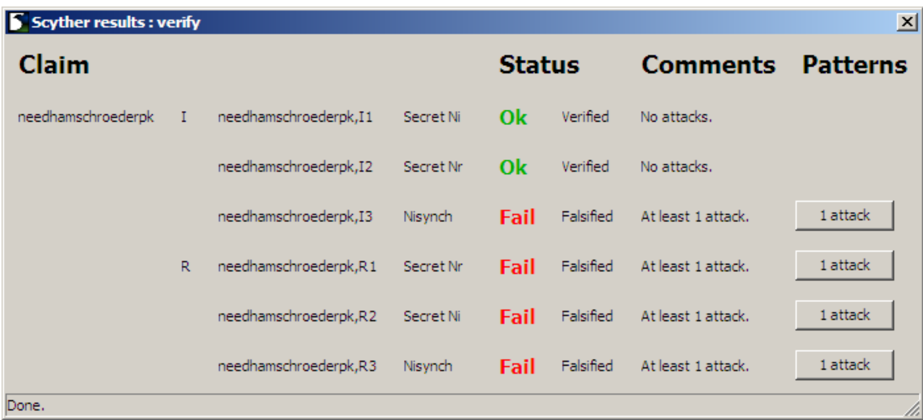


Fig. 4. Scyther: verification results

Often, the program finds does not find the attack and ends with the result "No attack within bounds”. In the verification process Scyther scans tree of all possibilities but is limited by the number of runs (program settings). It is important to choose an appropriate value. Too large value means more time needed for verification. A small amount may case that the attack is not found.

If you choose to show the attack found or any of the roles, the new window is opened. The main elements are rectangles connected by arrows. Arrows represent the order. Rectangle shows the new run, run event and the event caused by the security condition. The vertical axis represents the running protocol (roles occurence).

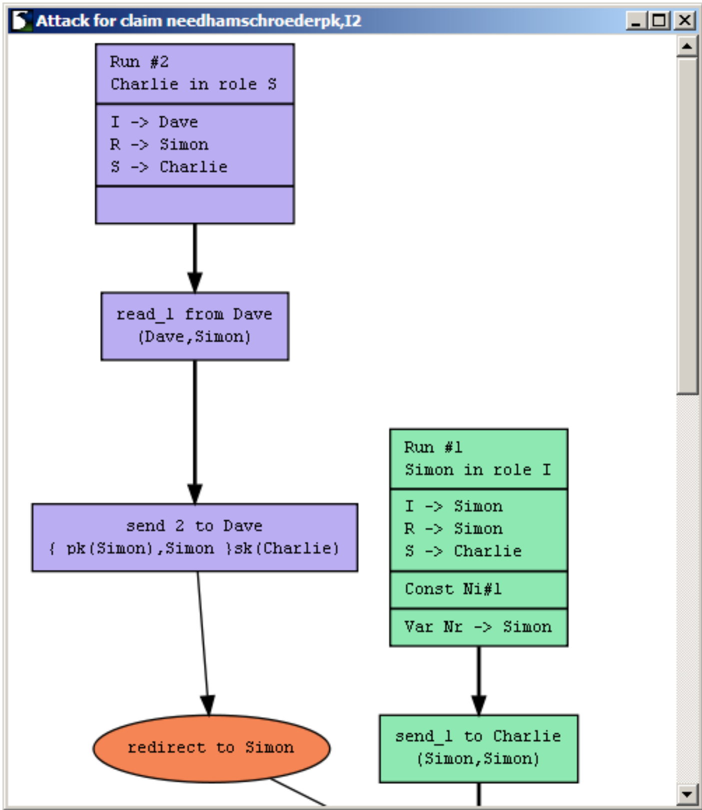


Fig. 5. Scyther: attack found

The resulted attack on Needham-Schroeder PK protocol found by the Scyther can be discussed with the attack presented in [3].

6 Conclusions

The aim of the paper was the presentation of Needham Schroeder specification in SPDL language. This language is the input for Scyther tool. This tool can be easily used for security protocol verification. The main advantage of the Scyther tool are the verification speed and the support of large number of communication connections. It can also verify the (sub)protocols in parallel. The main disadvantage is that we cannot change the settings within the protocol run and it is necessary to compile the protocol before we start the analysis.

Acknowledgements. This study was supported by the Ministry of Education, Youth and Sports of the Czech Republic (MSM0021630528 and MSM0021630503) and projects FIT-S-11-1, FIT-S-11-2 and FEKT/FIT-S-11-2.

References

1. Cremers, C.J.F.: The scyther tool: Verification, falsification, and analysis of security protocols. In: Gupta, A., Malik, S. (eds.) CAV 2008. LNCS, vol. 5123, pp. 414–418. Springer, Heidelberg (2008)
2. Needham, R., Schroeder, M.: Using Encryption for Authentication in Large Networks of Computers. CACM 21/12 (December 1978)
3. Lowe, G.: An Attack on the Needham-Schroeder Public-Key Authentication Protocol. Information Processing Letters (56), 131–133 (1995)
4. Mikus, P.: Tools for Environment for the Simulation of Communication, Master's thesis, FIT Brno University of Technology, Brno, Supervisor: Pavel Ocenasek (2010)

Routing Functionality in the Logic Approach for Authentication Protocol Design

Ocenasek Pavel and Hranac Jakub

Faculty of Information Technology,
Brno University of Technology,
Božetěchova 2, 612 66 Brno, Czech Republic
ocenaspa@fit.vutbr.cz, xhrana00@stud.fit.vutbr.cz
http://www.fit.vutbr.cz/~ocenaspa/

Abstract. This paper describes one of possible implementations of method used for protocol design described in a paper from 1998 by L. Buttyán, S. Staamann and U. Wilhelm which proposes a backward search (regression) when synthesizing an authentication protocol. Furthermore, the approach presented in this paper allows a designer to define participants who are trustworthy enough to transfer information between other two participants without existence of a direct channel to achieve basic routing functionality.

Keywords: Logic, Authentication, Key Distribution, Security Protocol, Routing.

1 Introduction

Logic which we are building on was first introduced in [1]. It serves as means for generating authentication protocols (a subset of security protocols). While creating automatic synthesis using this logic as its basis a few unexpected problems were discovered. Some could be solved by extending the set of heuristic rules to cover them (freshness of formulas for example) but one had to be solved by extending the original logic.

The paper covers introduction information, which can be found in details in [3]. Furthermore, it presents the partial results of the research published within the thesis [4].

First part of the paper describes the process of an automated protocol synthesis based on designer specified inputs. Second part introduces an idea of routing messages (allowing principals to act as 'routers' and deliver messages even when there is no direct or public channel connecting the source and the destination).

2 Message Ownership and Routing Extension

When implementing and testing the solution, one problem with message delivery was discovered. Core of the problem lays in the 'Seeing rule' defined in [1] which states:

If a principal P receives a message X via a channel C, and can read this channel, then recognizes that the message has arrived on C and P can see the message.

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 366–373, 2011.

$$\frac{P \triangleleft C(X), P \in r(C)}{P \mid\equiv (P \triangleleft X \mid C), P \triangleleft X} \tag{1}$$

While on the first look this rule may seem correct and sufficient, it does not say anything about someone being able to write the channel C. Imagine the following situation. There is a net of principals as described in Fig. 1.

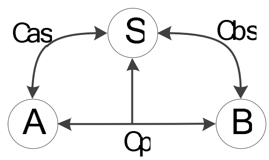


Fig. 1. Principals and channels connecting them each with S and one public channel Cp.

Now, the problem lies in composing communication steps for the goal:

$$A \mid\equiv B \mid\mid\sim A, Na \tag{2}$$

There is a set of assumptions and heuristic rules:

Heuristic rules

$$\begin{aligned} & P \triangleleft X \\ & \mapsto P \triangleleft C(X) \\ & \mapsto P \in r(C) \end{aligned} \tag{H1}$$

$$\begin{aligned} & P \mid\equiv Q \mid\mid\sim X \\ & \mapsto P \mid\equiv Q \mid\sim X \\ & \mapsto P \mid\equiv \#(X) \end{aligned} \tag{H2}$$

$$\begin{aligned} & P \mid\equiv Q \mid\sim X \\ & \mapsto P \triangleleft C(X) \\ & \mapsto P \in r(C) \\ & \mapsto P \mid\equiv (w(C) = \{Q\}) \mid P \mid\equiv (w(C) = \{P, Q\}) \\ & \mapsto Q \triangleleft X \end{aligned} \tag{H3}$$

$$\begin{aligned} & P \mid\equiv Q \mid\sim \alpha \\ & \mapsto P \triangleleft C(\alpha) \\ & \mapsto P \in r(C) \\ & \mapsto P \mid\equiv (w(C) = \{Q\}) \mid P \mid\equiv (w(C) = \{P, Q\}) \\ & \mapsto Q \mid\equiv \alpha \\ & \mapsto P \mid\equiv ((Q \mid\mid\sim \alpha) \rightarrow (Q \mid\equiv \alpha)) \end{aligned} \tag{H4}$$

$$\begin{aligned}&P \mid\equiv \alpha \\ &\mapsto P \mid\equiv \beta \\ &\mapsto P \mid\equiv \beta \to \alpha\end{aligned} \tag{H5}$$

Assumptions

$$\begin{aligned}&A \mid\equiv (w(Cas) = \{A, S\}); S \mid\equiv (w(Cas) = \{A, S\}); \\ &B \mid\equiv (w(Cbs) = \{B, S\}); S \mid\equiv (w(Cbs) = \{B, S\}); \\ &S \in r(Cas); A \in r(Cas); S \in r(Cbs); B \in r(Cbs); \\ &A \in r(Cp); A \in w(Cp); B \in r(Cp); B \in w(Cp); \\ &S \in r(Cp); S \in w(Cp); A \mid\equiv \#(Na); \\ &A \mid\equiv ((S \mid\mid\sim \alpha) \to (S \mid\equiv \alpha)); \\ &A \mid\equiv ((S \mid\equiv (B \mid\sim X)) \to (B \mid\sim X));\end{aligned} \tag{3}$$

By applying those we get to the point where we have to get done the goal:

$$B \lhd A, Na \tag{4}$$

This goal can be completed by applying the heuristic rule (H1) (introduced in [1] and based on (1)). At this point we generate one protocol step and we need to have defined which channels B can actually read. If it can read both (Cbs and Cp) then the correct (read 'correct in the logic') answer is to get two variants as solution for this goal:

$$\begin{aligned}&B \lhd Cp(A, Na) \\ &B \lhd Cbs(A, Na)\end{aligned} \tag{5}$$

The first one is correct, however the second one can be correct only if S (who is the only one who can write to Cbs except for B itself) already knows (A, Na). If this is not the case, then we have successfully generated a faulty protocol.

Even worse, when the channel Cp is removed then only the second variant from (6) gets generated and the protocol will never work.

First goal will be to eliminate generating of invalid rules. To solve this problem we introduced an attributed called 'message ownership'.

At the beginning we assign this attribute to principals who 'own' a fixed message. That means they know this message from the start of the communication. Now, it is important to own only fixed messages, because those are substituted strictly to 'itself' during synthesis while template messages can be any message at all (it is the same problem as with freshness test on nonces. If a nonce would not be a fixed message, principal would accept any message as fresh).

To use the owner attribute it is necessary to modify the 'Seeing rule' as follows:

$$\frac{P \lhd C(X), P \in r(C), Q \in w(C), Q \in owns(X)}{P \mid\equiv (P \lhd X \mid C), P \lhd X} \tag{6}$$

Meaning: P can see message X if it receives it via a channel it can read and there is someone else who can write to this channel and knows (owns) the message.

This modification of the logic removes problem with expecting messages on channels where no principal knows them. It works well in 'routed' environment where there always exists a direct[1] channel (at least insecure channel) between two principals.

3 Routing

When there is basis of message ownership in the logic it is possible to build on this and set up a basic routing for messages[2]. Imagine the situation form Fig. 1 when the common channel (Cp) gets removed and the goal stays the same. There still exists a route for message (A, Na) however it must be retransmitted by the principal S.

Plus, when there would be more principals like S (who connect A and B) it may be useful to allow protocol designer to specify which principals can send what messages, from which sources to which principals. This defines a basic trust levels between principals.

Note: Only messages that would normally be transmitted via insecure channels should be allowed to be retransmitted by any principals.

To cover this functionality we defined three new attributes. CanResend and canResendFrom when each of those takes a message (fixed or template) as one parameter and optionally a principal as other parameter. The last attribute is canOwn which is a meta-attribute. It must not be set in assumptions and is only used in heuristic rules. It encapsulates routing rights.

Heuristic rules need to be modified to reflect those changes, so from (H1) we get those new rules:

$$\begin{aligned} & P \lhd X \\ & \mapsto P \lhd C(X) \\ & \mapsto P \in r(C) \\ & \mapsto Q \in w(C) \\ & \mapsto Q \in canOwn(X) \mid Q \in owns(X) \end{aligned} \tag{H1A}$$

$$\begin{aligned} & P \in canOwn(X) \\ & \mapsto P \lhd C(X) \\ & \mapsto P \in r(C) \\ & \mapsto Q \in w(C) \\ & \mapsto P \in canResend(X) \mid P \in canResendFrom(X, Q) \\ & \mapsto Q \in canOwn(X) \mid Q \in owns(X) \end{aligned} \tag{H1B}$$

---

[1] In terms of this logic all channels are direct. It does not say anything about how is the channel physically realized.

[2] Note that routing is never used for formulas because only a formula principal believes in can be transmitted which means principal either builds its own formula or it's already received from somewhere else.

Note: Notice that trust is set in a backwards manner. It is not set on route from source (setting whom we entrust with a message) but rather from destination to source (setting who is reliable source). Power of this reasoning stays the same.
Note: It is not important to worry about cycles since those are already tested for during cycle detection. So message will never travel through same principal twice. In complex networks where everyone is connected with everyone except for two principals who wish to exchange a message there will be as many protocols as there are routes in an oriented graph where nodes (principals) are connected by using channels.

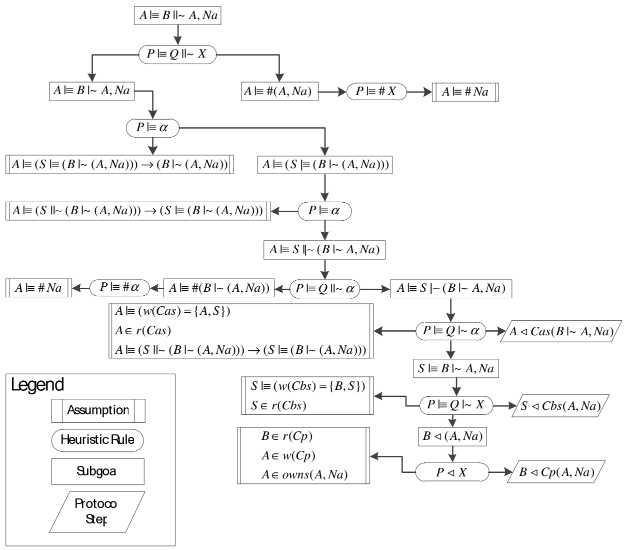


Fig. 2. Synthesis output

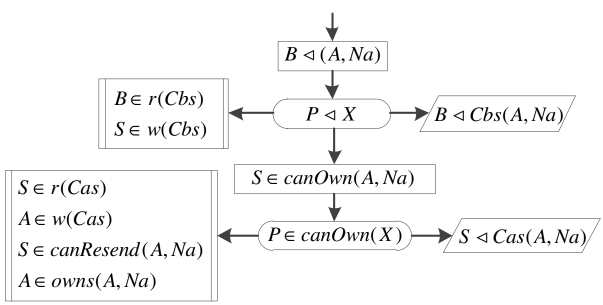


Fig. 3. Modified part of synthesis with routing enabled

Let's see how this will work by demonstrating functionality on the goal (2) and net defined by Fig. 1 using assumptions (3) and rules (H2-H5, H1A, H1B). First, nobody has the canResend attribute set (see Fig. 2.).

Synthesis has successfully filtered out problematic protocol step $B \lhd Cbs(A, Na)$. Next demonstration will show how routing can be done by giving S the attribute *canResend* for the message (A, Na). To reduce amount of found protocols we remove the channel Cp. Only part from $B \lhd (A, Na)$ gets extended as shown in Fig. 3.

The message gets properly routed through S and reproduced to B so it can then send it back. Although this example is not very complex, it demonstrates well how routing works. Same can be done with multiple principals routing messages, however amount of generated protocols gets higher with each possible route from A to B in the principals net.

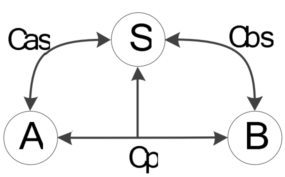
4 Benchmarking

Message owner modification only adds one attribute which should not affect performance too much. Routing can increase synthesis time especially with complex net of principals where there are many routes for a message to travel between two principals. Following test uses rules (H2-H5, H1A, H1B) and always has one goal:

$$A \mid\equiv B \parallel\!\sim A, Na \tag{7}$$

Set of assumptions is defined by a principal net (for channel read/write permissions, channels connecting two principals only are considered secure). Message resend attribute is noted in each case separately as well as any other assumptions which are not channel related.

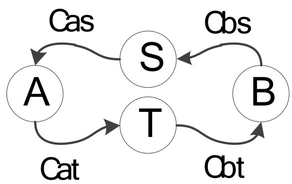
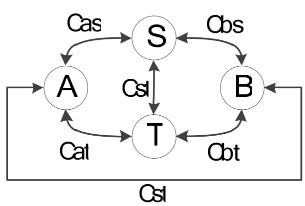
Table 1. Performance with and without using routing extension[3]

Principal net scheme	Results for message owner extension	Results for message routing extension[4]
 $S \in canResend(A, Na)$; $A \mid\equiv \#(Na)$; $A \mid\equiv ((S \parallel\!\sim \alpha) \rightarrow (S \mid\equiv \alpha))$; $A \mid\equiv ((S \mid\equiv (B \mid\!\sim X)) \rightarrow (B \mid\!\sim X$	Results: 1 Average time: 41ms Solution: $B \lhd Cp(A, Na)$ $S \lhd Cbs(A, Na)$ $A \lhd Cas(B \mid\!\sim A, Na)$	Results: 1+4 Average time: 76ms One of solutions: $S \lhd Cas(A, Na)$ $B \lhd Cbs(A, Na)$ $S \lhd Cbs(A, Na)$ $A \lhd Cas(B \mid\!\sim A, Na)$

3 Tested on Intel® Core™ 2 Duo @ 2.66GHz, 64-bit OS; time is an average time calculated from 1000 tries.

4 Results for routing extension always include the original ones as well so the amount of available communication schemes is numbered as x+y where x is the number of original solutions and y amount of new ones.

Table 1. (*continued*)

 $S,T \in canResend(A,Na)$; $A \mid\equiv \#(Na)$; $A \mid\equiv ((S \Vert\!\sim \alpha) \rightarrow (S \mid\equiv \alpha))$; $A \mid\equiv ((S \mid\equiv (B \mid\!\sim X)) \rightarrow (B \mid\!\sim X))$	Results: 0 Average time: 30ms No solution was found.	Results: 0+1 Average time: 61ms Solution: $T \lhd Cat(A,Na)$ $B \lhd Cbt(A,Na)$ $S \lhd Cbs(A,Na)$ $A \lhd Cas(B \mid\!\sim A,Na)$
 $S,T \in canResend(A,Na)$; $A \mid\equiv \#(Na)$;	Results: 1 Average time: 55ms Solution: $B \lhd Cab(A,Na)$ $A \lhd Cab(A,Na)$	Results: 1+4 Average time: 160ms One of solutions: $T \lhd Cat(A,Na)$ $S \lhd Cst(A,Na)$ $B \lhd Cbs(A,Na)$ $A \lhd Cab(A,Na)$

5 Conclusions

While routing has almost double increase in resources needed to find valid protocol, it has a chance of creating several new correct protocols. However it is necessary to avoid assigning canResend permission to everyone (use only reliable principals and those which you actually need). Dense principal nets make synthesis to take much longer to complete when routing is allowed to multiple nodes and synthesis returns rather large amount of possible protocols.

Acknowledgements. This study was supported by the Ministry of Education, Youth and Sports of the Czech Republic (MSM0021630528 and MSM0021630503) and projects FIT-S-11-1, FIT-S-11-2 and FEKT/FIT-S-11-2.

References

1. Buttyán, L., Staamann, S., Wilhelm, U.: A Simple Logic for Authentication Protocol Design, In: 11th IEEE Computer Security Foundations Workshop (1998)
2. Zhou, H., Foley, S.N.: Fast automatic synthesis of security protocols using backward search. In: Proceedings of the 2003 ACM Workshop on Formal Methods in Security Engineering, Washington, D.C., October 30 (2003)
3. Ocenasek, P., Hranac, J.: Regression Based Logic for Authentication Protocol Design. In: Proceedings of 2010 International Conference on Communication and Vehicular Technology. IEEE, Chengdu (2010)
4. Hranac, J.: Methods of the Security Protocols Design, Master's thesis, FIT Brno University of Technology, Brno, Supervisor: Pavel Ocenasek (2010)

An Approach for Security Protocol Design Based on Zero-Knowledge Primitives Composition

Očenášek Pavel

Faculty of Information Technology,
Brno University of Technology,
Božetěchova 2, 612 66 Brno, Czech Republic
ocenaspa@fit.vutbr.cz
http://www.fit.vutbr.cz/~ocenaspa/

Abstract. The paper deals with automated methods for the design of security protocols and their design using zero knowledge protocols, or protocols, where it is possible to use zero knowledge protocols such as subprotocols.. Specific emphasis is placed on the use of compositional method. The paper also include the example of protocol design.

Keywords: Zero knowledge, Security Protocol, Automation, Implementation.

1 Introduction

Zero knowledge protocols [4] [5] have been created because the traditional security protocols for leads to information disclosure in any running protocol, which if they accumulate sufficient quantities can be used for example for identity theft. When running Zero Knowledge (we ale use the abbreviation ZK) protocol does not log when running leak any information.

This paper presents the partial results of the research published within the thesis [1].

When running zero knowledge protocol, participant is trying to prove knowledge of certain information to the other party. For participants we use the abbreviation P (prover) for those who are trying to show that knows the information. We use the abbreviation V (verifier) for the participants which are trying to verify this knowledge. The main idea is the ability to carry proof of knowledge of information without learning anything other than that the evidence is true. This is not about proof in a mathematical sense, but it is a dynamic process between the two parties.

ZK protocols use for maintaining the secrecy such problems as factoring large prime numbers and the Knapsack problem. Compared to modern cryptography, ZK protocol are less time efficient, i.e. the run takes longer, which can be disadvantageous for RT applications, on the other hand ZK are not as computationally intensive and can be used such as embedded devices, Smart cards, etc.

The disadvantages of ZK are their usefulness, because it transmits no information at runtime, they cannot be used for transmission of messages or key distribution. But can be used for authenticating users and there are cases where it can also be used to

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 374–378, 2011.

confirm the validity of data entered, without the recipient party, which requires confirmation of the data he knew.

2 Zero Knowledge Protocol Design

Traditional notation of Feig-Fiat-Shamir [3] can generally be used for any other ZK protocol notation. ZK protocol will always have the stated sequence of messages: commitment, challenge, response, and final confirmation or rejection.

The main drawbacks of the proposal of separate ZK protocol thus lies rather in the creation of mathematical functions, which are possible to use in this type of protocols. This work would not be used as part of the proposal to address this ZK but rather deals with the question og how the method can be used for constructing more complex protocols, where ZK protocols are used as subprotocols for the identification of individual participants.

2.1 Input and Output Knowledges

The requirements for ZK protocol can be specified as follows:

- Input knowledge

$A = \{Ncommit, f(x,y)\}$
$B = \{Naccept, Nchallenge\}$

- output knowledge

$A = \{Nchallenge, Naccept\}$
$B = \{Ncommit, f(Ncommit, Nchallenge)\}$

Unfortunately, these requirements can not express order of messages, that B has to send *Naccept* after receiving *f(Ncommit, Nchallenge)* from the party A. We have to specify this information in requirements, or include the information that proper subject have to authenticate using ZK protocols.

3 Implementation

The program seeks to gradually meet the specified goals. As the primary goals, the selected objectives are defined in the protocol specification. The program seeks to gradually meet the objectives with the use of some primitive. If there is a primitive, which fulfills the goals, its assumptions become the part of new goals. Upon completion of all the assumptions, the primitive is added to the protocol and knowledge are added to the Prolog database, as defined in the primitive specification.

The program continues this way until all the goals are met. Primitives, from which the program selects appropriate, are defined in the program startup. All primitives are nondestructive, therefore it is not necessary to examine whether the addition of a primitive changes any assumptions on which it depended on some of the previous primitives.

There are other additional conditions placed for the primitives in the design process of the protocol. None of the constraints may not be in any state of the current protocol. If the primitive is found to satisfy a goal, but its implementation becomes the restrictive conditions, it is necessary to choose another primitive.

Primitives themselves must also logically follow. For this purpose, we use lists of preactive and postactive users. Preactive user of each primitive must also belong to the set of postactive users from previous primitive. In the case that is the first primitive of the protocol, the preactive user must be defined as an initiator of the communication in the protocol specification. If the conditions are not met, the protocol is not valid.

The program has no special interface and takes advantage of Prolog interpreter.

4 Protocol Example: Key Distribution

The presented example shows the protocol for distribution of the generated key between both participating subjects *a* and *b*, with the trusted server *s*.

```
actor(a).
actor(b).
actor(s).
server(s).
channel(a, b).
channel(b, a).
channel(a, s).
channel(s, a).
channel(b, s).
channel(s, b).
key(a, b, secret, 0).
key(a, s, secret, 0).
key(b, s, secret, 0).
knows(s, key(a, b, secret, 0)).
knows(s, key(a, s, secret, 0)).
knows(s, key(b, s, secret, 0)).
knows(a, key(a, s, secret, 0)).
knows(b, key(b, s, secret, 0)).
initiator(a).
```

```
goal(knows(a, key(a, b, secret, 0))).
goal(knows(b, key(a, b, secret, 0))).
forbidden(knows(s, x)).
forbidden(knows(spy, x)).
forbidden(knows(spy, key(a, b, secret, 0))).
forbidden(knows(spy, key(a, s, secret, 0))).
forbidden(knows(spy, key(b, s, secret, 0))).
```

The generated protocol is as follows:

```
1: a -> s: s, {|req(a, reqI(0, Kab0)), a, b, s|}Kas0,
             <req(a, reqI(0, Kab0)), a, b, s>Kas0
2: s -> a: a, <req(a, reqI(0, Kab0)), a, b, s>
3: a -> b: b, reqI(0, Kab0), a, b, s, <reqI(0, Kab0),
             a, b, s>
4: b -> a: a, <reqI(0, Kab0), a, b, s>
5: b -> s: s, {|req(b, reqI(0, Kab0)), a, b, s|}Kbs0,
            <req(b, reqI(0, Kab0)), a, b, s>Kbs0
6: s -> b: b, <req(b, reqI(0, Kab0)), a, b, s>
7: s -> a: a, {|Kab0|}Kas0, <req(a, reqI(0, Kab0)),
           Kab0, a, b, s>Kas0
8: s -> b: b, {|Kab0|}Kbs0, <req(b, reqI(0, Kab0)),
           Kab0, a, b, s>Kbs0
```

5 Conclusions

The presented approach implements the composition method presented in [2]. As we have showed in the example, the application is able to design new simple protocol upon the basic requirements and goals specification. The protocols can then include the zero knowledge primitives, as the basic composition primitives.

Acknowledgements. This study was supported by the Ministry of Education, Youth and Sports of the Czech Republic (MSM0021630528 and MSM0021630503) and projects FIT-S-11-1, FIT-S-11-2 and FEKT/FIT-S-11-2.

References

1. Safar, J.: Zero-knowledge protocol design, Master's thesis, FIT Brno University of Technology, Brno, Supervisor: Pavel Ocenasek (2010)
2. Choi, H.-J.: Security protocol design by composition. Technical report, University of Cambridge (2006)

3. Fiege, U., Fiat, A., Shamir, A.: Zero knowledge proofs of identity. In: STOC 1987: Proceedings of the Nineteenth Annual ACM Symposium on Theory of Computing, pp. s.210–s.217. ACM, New York (1987) ISBN 0-89791-221-7
4. Rosen, A., Goldreich, O.: Concurrent Zero-Knowledge. Springer, Heidelberg (2006) ISBN 3-540-32938-2
5. Simari, G.I.: A Primer on Zero Knowledge Protocols (2002)

Part V

Touch and Gesture Interfaces

Effects of Joint Acceleration on Rod's Length Perception by Dynamic Touch

Takafumi Asao[1], Yuta Kumazaki[2], and Kentaro Kotani[1]

[1] Faculty of Engineering Science, Kansai University
[2] Graduate School of Science and Engineering, Kansai University,
3-3-35 Yamate-cho, Suita, Osaka 564-8680, Japan
{asao,k738600,kotani}@kansai-u.ac.jp

Abstract. Studies on dynamic touch have indicated that humans can estimate the length of a rod held in one hand simply by wielding it, without any visual information. Traditionally, these types of studies have held that proprioception is important for perceiving the moment of inertia of the rod, but this has not been demonstrated experimentally. In the present paper, we focused on joint acceleration and torque as physical values that approximate proprioception by introducing dynamics-based mechanics. In the experiments, the acceleration of the swinging of the rod was controlled. Since the acceleration and the torque both varied with time, their peak values were adopted as representative values. No correlation was observed between the angular acceleration and the perceived rod length. However, a strong correlation was found between joint torque and the perceived length. This indicates that humans perceive rod length through joint torque, which approximates proprioception.

Keywords: Dynamic touch, joint acceleration, joint torque, proprioception.

1 Introduction

The technologies of virtual reality are presently quite developed for visual and auditory information. In addition, haptic interfaces are now being developed that can provide tactile and kinesthetic information. One of these kinesthetic devices presents a virtual rod of a certain length [1]. The basis of human perceptual for haptic and kinesthetic information needs to be understood for effective development of these types of haptic interfaces.

Humans can perceive the length of a rod by simply grasping or wielding it, in the absence of visual information. This perceptual subsystem is called *dynamic touch* [2][3], and it functions through a combination of proprioceptive information and cutaneous tactile sensation. Current studies on dynamic touch indicate that humans might perceive rod length from the invariant inertial moment of the rod [3][4]. Solomon & Turvey [5] conducted experiments in which the frequency of a swinging rod was manipulated by confining the spatial amplitude of the rod's swing, and they found that the frequency did not affect the length perception. Stephen & Arzamarski [6] determined that the accuracy of perceptual length was progressive, as participants

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 381–390, 2011.

were able to determine the rod's length from vibratory information created by striking the rod. These studies offered important knowledge about haptic perception, but they also indicated physical invariants affecting length perception, on the basis of statistical analysis. Although many researchers have commented that proprioception is important for dynamic touch based on preliminary studies of dynamic touch [2][3][7], none of these studies verified the particular role of proprioception.

Proprioception that arises from dynamic touch can be defined by a motion equation that describes the wielding of a rod. Fig. 1 shows a link model for wielding a rod by the wrist joint. For simplicity, the wrist pin joint is also the rod joint. The motion equation is then expressed by following equation:

$$\tau = \left(I + ml_g^{\ 2}\right)\ddot{\theta} + mgl_g \cos\theta , \tag{1}$$

where τ is the joint torque, m is the mass of the rod, I is the moment of inertia of the rod about its center of mass (CM), g is acceleration due to gravity, l_g is the distance between the CM of the rod and the wrist joint, and θ is the wrist joint angle. As indicated on the right side of the equation, the joint torque τ consists of an inertia term and a gravity term. Because the joint torque is generated by muscles, the human can indirectly perceive the joint torque from the proprioception in the muscles. Therefore, it should be easy to perceive the moment of inertia I when the joint acceleration $\ddot{\theta}$ is large. Kingma & Beek [5] remarked that humans may perceive length by using the moment of inertia I when the joint acceleration is large, and by using static torque, which is the second term in the equation (1), when the joint acceleration is small, similar to static holding. Their discussion therefore supports our idea. In the present paper, we focus on the joint acceleration as a close approximation to proprioception; it clarifies the relationship between the wrist joint acceleration and the rod's perceived length.

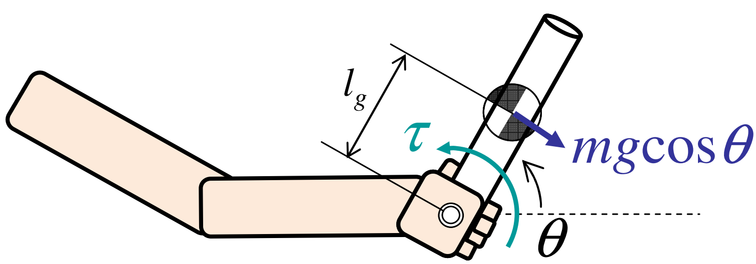


Fig. 1. Simple model of wrist flexion while holding a rod

2 Experiment

Two methods were used to control the acceleration of the bending wrist joint. One method was to change the swing frequency; participants had to wield the rod in rhythm to a metronome at 0.5 or 1.0 Hz. The second method was to change the wielding mode so that the rod was wielded either *softly* or *strongly*.

2.1 Materials

The rods used in the present study were 20 mm in diameter and 400, 500, or 600 mm long and were made of either wood (density 0.66 g/cm^3) or aluminum (2.7 g/cm^3). Sponge grips were attached to the lower portions of the rods to prevent effects of tactile feeling and specific heat of the rods on the length perception. The internal and external diameters of the grips were 20 and 27 mm, respectively. The mass of each grip was 23 g.

2.2 Apparatus

The right arm of each participant was restrained by hook and loop fasteners on an armrest, as shown in Fig. 2, to avoid the influence of the proprioception of elbow and shoulder joint angles on the length perception. The tightness of the restraints was such that the arm was not put to sleep. This restraint allowed participants to swing the rods only by using radial and ulna flexion. A goniometer (M110, P&G) was attached to the right wrist on the opisthenar side for measuring the wrist angle. The radioulnar wrist angle was recorded by a multi telemeter (Web-9000, Nihon Koden) at a sampling frequently of 200 Hz.

Fig. 3 shows the apparatus used to express the perceived length. Participants could move a cursor that was clamped to a rubber belt by pulling on the belt with their left hands, since the belt was a pulley block. Therefore, participants could indicate the perceived tip of each rod using this apparatus. The distance between the cursor position and the wrist position was measured as the perceived length of the rod.

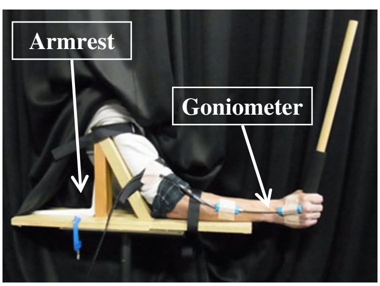


Fig. 2. Setup for the right arm

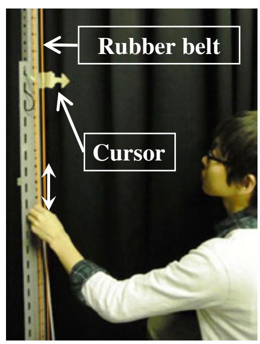


Fig. 3. Apparatus for indicating perceived length

2.3 Procedure

Under the *strongly wielded condition*, participants were instructed that they should swing strongly in order to achieve a high angular speed and acceleration for their wrist joint angle. Under the *softly wielded condition*, they were instructed to swing softly and smoothly so that no impulsion would be felt when they swung the rod down. Before the experiments, participants practiced swinging the rods to comply with these swinging instructions.

In the experiments, each participant first sat in a chair and with the right forearm and upper arm on the armrest. The goniometer was threaded onto the right wrist on the opisthenar side. The right arm was hidden by a curtain to prevent the participant from seeing the right arm or the rods. The frequency of a metronome was set at 0.5 or 1.0 Hz, and the rhythm was started. One rod was then selected randomly and thrust into the right palm. The participant grasped the rod firmly so that the rod would not shake in their palm when it was wielded. The participant then swung the rod *softly* or *strongly* with the metronome rhythm; they swung the rod downward every time the metronome ticked. Each trial was conducted in a completely randomized order and the trials were replicated 6 times for each condition. The entire experiment took three days for each participant, and 48 trials per day were conducted. Participant fatigue was prevented by allowing 5 minutes breaks after every 12 trials. A summary of the experimental conditions is shown in Table 1.

2.4 Participants

Participants were five male university students. They were informed about the experiments beforehand and consented to participate.

Table 1. Experimental conditions

Participants	5 male university students	
Wielding conditions	strong, soft	
Frequency of swing Hz	0.5, 1.0	
Materials (density g/cm^3)	wood (0.66), aluminum (2.7)	Sponge grip
Length mm	400, 500, 600	215
Outer diameter (Inner diameter) mm	20	27 (20)
Order of trials	completely random	

3 Results

Fig. 4 shows the wrist angle and angular acceleration for the softly wielded condition under swinging frequency of 0.5 and 1.0 Hz, respectively. Similarly, Fig. 5 shows the angle and acceleration for the strongly wielded condition. In the figures, a vertical positive value means radial flexion and a negative value denotes ulnar flexion. The acceleration was calculated by 2nd order centered difference. As shown in the figures, the amplitudes of the wrist angles were not very different. The accelerations had positive peak values when the wrist joint was most inflected to the ulnar side. The peak values were larger for swinging frequency of 1.0 Hz than for 0.5 Hz. A comparison of the perceived length with the wrist angle and the angular acceleration was difficult, since these varied with time. According to equation (1), the wrist joint torque should become larger with an increase in the angular acceleration. If the torque is large, the quantity of proprioceptive information should also be large. Therefore, the mean peak value of the acceleration was employed as the representative value of the angular acceleration (RVAA). Two-factor analysis of variance (soft or strong swinging and the frequency 0.5 or 1.0 Hz of the metronome) was performed on the mean RVAA.

This revealed main effects of the swinging manner ($F(1, 716) = 487.81$, $p < 0.001$) and the frequency ($F(1, 716) = 20.99$, $p < 0.001$). The interaction was not significant ($F(1, 716) = 17.84$, $p = 0.35$). In the following analysis, RVAA was then substituted for the swinging manner and the acceleration.

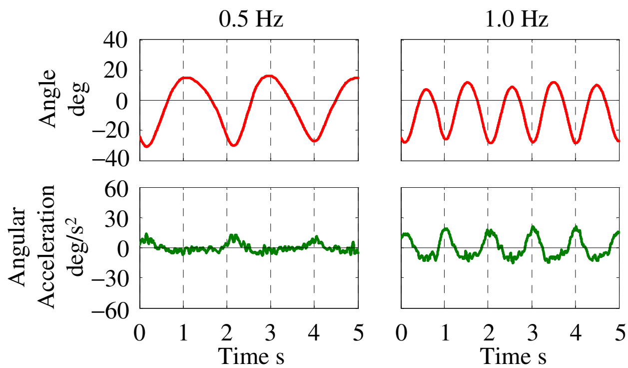


Fig. 4. Wrist angle and angular acceleration for the softly wielded condition

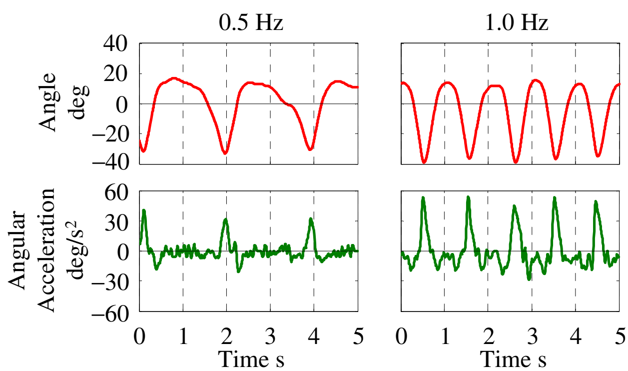


Fig. 5. Wrist angle and angular acceleration for the strongly wielded condition

The relationship between the subject means of RVAA and that of the perceived length is shown in Fig. 6. Smaller RVAAs were found for the softly wielded condition (triangle plots) than for the strongly wielded condition (circle plots). No correlation was noted between RVAA and perceived length, since the contribution ratio was 0.04. However, the contribution ratio for the strongly wielded condition alone was 0.70, which reflected the fact that swinging at high acceleration was easy for a lightweight rod (dark color plots denote the lighter wooden rods).

Fig. 7 shows the relationship between the moment of inertia of the rods and the perceived length. The contribution ratio was 0.96. Similar to the present results, a high correlation between these factors has often been reported in other traditional dynamic touch studies [3][4][5][6], which concluded that humans could estimate a rod length by perceiving the moment of inertia.

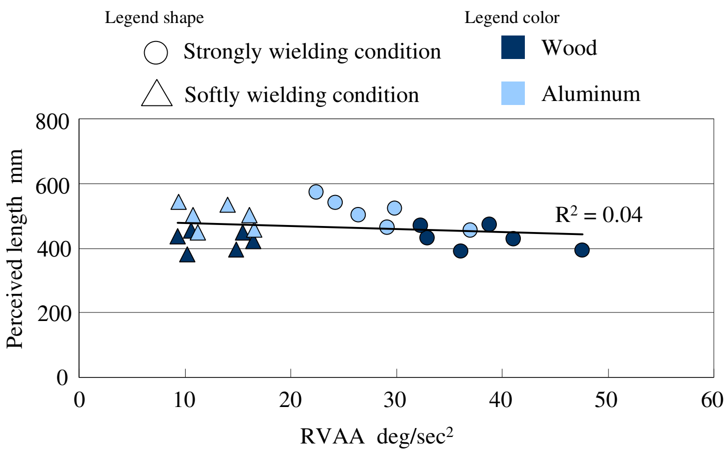


Fig. 6. Relationship between RVAA and perceived length

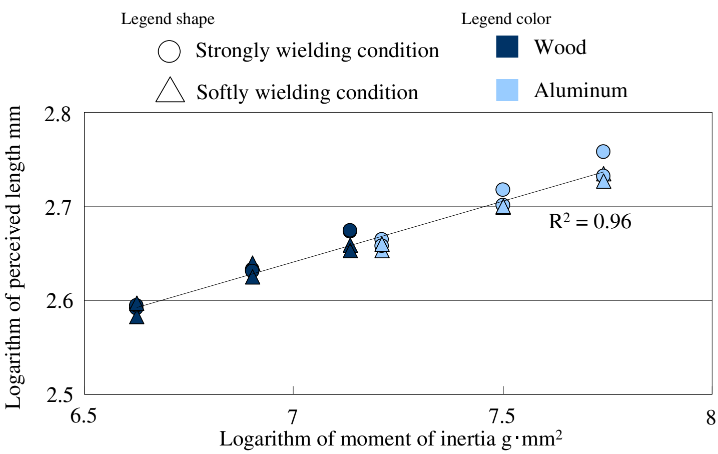


Fig. 7. Relationship between moment of inertia and perceived length

4 Discussion

As shown in Fig. 6, no correlation was apparent between RVAA and the perceived length. On the other hand, the wrist joint torque became larger with increases in the angular acceleration, as predicted by equation (1). Therefore, the representative value of the joint torque (RVJT) corresponding to RVAA was calculated (see appendix for details). Fig. 8 shows the relationship between the mean RVJT and the mean perceived length with their contribution ratio. The correlation was strong, with a contribution ratio of 0.82 between RVJT and the perceived length. Accordingly, humans may estimate the rod length by perceiving the peak joint torque as a way of perceiving the rotational inertia by proprioception. Humans may also use other information for length perception, since the correlation between RVJT and the perceived length was not as strong as that between the moment of inertia and the perceived length. Because a correlation was found between the angular acceleration and the perceived length for the strongly wielded condition in Fig. 6, the possibility existed that the angular acceleration was also one of the clues used for perceiving the length, which could also be perceived by proprioception.

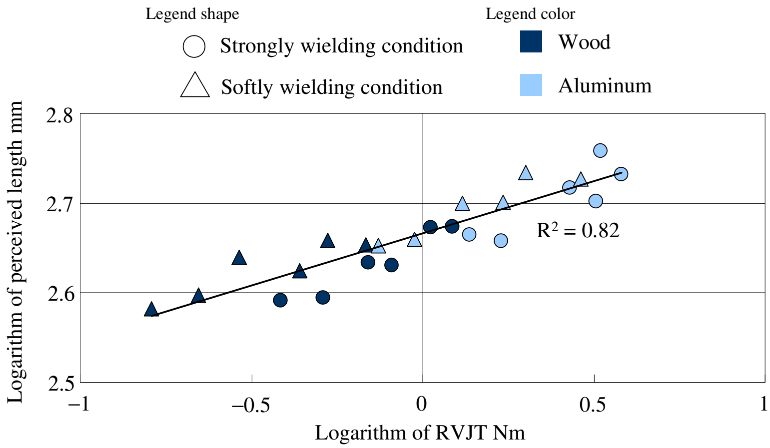


Fig. 8. Relationship between RVJT and perceived length

5 Conclusions

In the present paper, we introduced dynamics-based mechanics to clarify the relationship between angular acceleration and joint torque, and also to clarify the perception of length. No correlation was found between RVAA and the perceived length, but a high correlation was found between RVJT and the perceived length. However, the contribution ratio between RVJT and the perceived length was lower than that between the moment of inertia and the perceived length.

In this study, the peak value of the angular acceleration and the joint torque were selected as representative values, because the acceleration and the torque varied with time. In fact, humans may also use other physical values for estimating length. Therefore, a methodology that will evaluate these time-varying values will be developed in future studies.

References

1. Ikeda, T., Matsuda, H., Nakamura, T., Shiota, Y., Sakamoto, K., Shimizu, Y.: A Method of Presenting Haptic Information for Dynamic Touch Using Superficial Moment of Inertia. Transactions of The Institute of Electronics, Information and Communication Engineers J89-D(6), 1403–1412 (2006)

2. Gibson, J.J.: The Senses Considered as Perceptual Systems, pp. 127–128. Houghton Mifflin, Boston (1966)
3. Turvey, M.T.: Dynamic Touch. American Psychologist 51(11), 1134–1152 (1996)
4. Kingma, I., Langenberg, R., Beek, P.J.: Which Mechanical Invariants Are Associated With the Perception of Length and Heaviness of a Nonvisible Handheld Rod? Testing the Inertia Tensor Hypothesis. Journal of Experimental Psychology: Human Perception and Performance 30, 346–354 (2004)
5. Solomon, H.Y., Turvey, M.T.: Haptically Perceiving the Distances Reachable With Hand-Held Objects. Journal of Experimental Psychology: Human Perception and Performance 14(3), 404–427 (1988)
6. Stephen, D.G., Arzamarski, R.: Self-training of Dynamic Touch: Striking Improves Judgement by Wielding. Attention, Perception, & Psychophysics 71(8), 1717–1723 (2009)
7. Loomis, J.M., Lederman, S.J.: Tactile Perception. In: Boff, K.R., Kaufman, L., Thomas, J.P. (eds.) Handbook of Perception and Human Performance, pp. 31.01–31.41 (1986)

Appendix: Calculating Theoretical Joint Torque

The moment of inertia I_1 of cylindrical rod on its center of mass CM_1 is described as

$$I_1 = \frac{m_1}{12} l_1^{\,2} + \frac{m_1}{4} r_1^{\,2}, \tag{1}$$

if the rod is homogeneous, where m_1, l_1, and r_1 are mass, length, and outer radius of the rod, respectively.

The moment of inertia I_2 of a grip on its center of mass CM_2 is

$$I_2 = \frac{m_2}{12} l_2^{\,2} + \frac{m_2}{4} \left(r_{2o}^{\,2} + r_{2i}^{\,2} \right), \tag{2}$$

where m_2, l_2, r_{2o}, and r_{2i} are mass, length, outer radius, and inner radius of the grip, respectively.

In the next step, the moment of inertia I_1' of a rod on a wrist joint O_{wrist} is developed. The left figure in Fig. 9 shows dimensions of the rod and the wrist. By the parallel axis theorem,

$$I_1' = I_1 + m_1 \left[\left(\frac{l_1}{2} - \frac{w_{wrist}}{2} \right)^2 + l_{w-g}^{\,2} \right], \tag{3}$$

where w_{wrist} is wrist breadth and l_{w-g} is length from wrist to center of grip.

In the same manner as for the rod, the moment of inertia I_2' of the grip on the wrist joint O_{wrist} (see also the right figure in Fig. 9) is

$$I_2' = I_2 + m_2 \left[\left(\frac{l_2}{2} - \frac{w_{wrist}}{2} \right)^2 + l_{w-g}^{\,2} \right]. \tag{4}$$

Therefore, the combined moment of inertia I is

$$I = I_1' + I_2' . \tag{5}$$

Joint torque τ around the wrist joint for swinging the rod can be calculated by

$$\tau = \left(I + m l_g^{\,2}\right)\ddot{\theta} + m g l_g \cos\theta , \tag{6}$$

where θ denotes the wrist angle from a horizontal plane, and combined mass $m = m_1 + m_2$. The length l_g from the wrist joint to the combined CM is

$$l_g = \frac{1}{2}\frac{m_1 l_1 + m_2 l_2}{m_1 + m_2} . \tag{7}$$

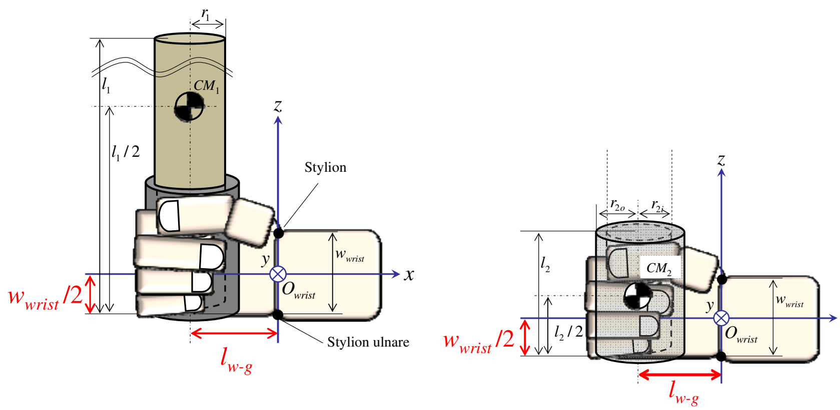


Fig. 9. Coordinate system and dimension with rod and grip

Appendix: Anthropometric Dimensions and Inertial Data of Rods

The mass and moment of inertia on the CM for each rod used in the experiments are shown in Table 2. Physical parameter of rods.

Anthropometric data and equivalent moment of inertia I, which was calculated by equation (6) for each participant, are shown in Table 3.

Table 2. Physical parameter of rods

Material	l_1 mm	m_1 g	I_1 kg m^2	I_2 kg m^2
Wood	400	83	1.11×10^{-3}	0.09×10^{-3}
	500	104	2.16×10^{-3}	
	600	124	3.74×10^{-3}	
Aluminum	400	339	4.53×10^{-3}	
	500	424	8.85×10^{-3}	
	600	509	15.3×10^{-3}	

Table 3. Equivalent rotational inertia for each participant

Participant	w_{wrist} mm	$l_{w\text{-}g}$ mm	Material	l_1 mm	I kg m^2
A	54	79	Wood	400	4.49×10^{-3}
				500	8.34×10^{-3}
				600	14.2×10^{-3}
			Aluminum	400	17.2×10^{-3}
				500	33.0×10^{-3}
				600	56.8×10^{-3}
B	48	76	Wood	400	4.54×10^{-3}
				500	8.44×10^{-3}
				600	14.3×10^{-3}
			Aluminum	400	17.4×10^{-3}
				500	33.3×10^{-3}
				600	57.4×10^{-3}
C	50	78	Wood	400	4.54×10^{-3}
				500	8.42×10^{-3}
				600	14.3×10^{-3}
			Aluminum	400	17.4×10^{-3}
				500	33.3×10^{-3}
				600	57.3×10^{-3}
D	50	79	Wood	400	4.56×10^{-3}
				500	8.45×10^{-3}
				600	14.3×10^{-3}
			Aluminum	400	17.4×10^{-3}
				500	33.4×10^{-3}
				600	57.3×10^{-3}
E	52	77	Wood	400	4.49×10^{-3}
				500	8.35×10^{-3}
				600	14.2×10^{-3}
			Aluminum	400	17.2×10^{-3}
				500	33.0×10^{-3}
				600	56.9×10^{-3}

ERACLE: Electromyography System for Gesture Interaction

Paolo Belluco, Monica Bordegoni, and Umberto Cugini

Dipartimento di Meccanica Politecnico di Milano,
Via La Masa 1, 20156 Milano, Italy
paolo.belluco@mecc.polimi.it,
{monica.bordegoni,umberto.cugini}@polimi.it

Abstract. Gesture interaction is one of the most important topics in the human-computer interaction. In this field, the main research activities are oriented on recognizing and classifying different gestures in order to interact with the computer directly with the body, without using classical mobile devices such as touchpad or trackball. This paper describes the development and the testing of our wearable interaction system that uses surface electromyography (sEMG) signals to recognize and process the gestures of the users. The core of the system is the "Eracle-board" that is a wearable 3-channel board developed in order to acquire the sEMG signals from the user's forearm. The acquired data are subsequently processed by an external device, which allows us to recognize and classify seven different gestures through the implementation of a neural network. Finally, the effectiveness of the system has been evaluated through some tests carried out with users.

1 Introduction

Acquiring and understanding the movement and the posture of the body is the new way to interact with computers. In fact, a large part of the research activities in human-computer interaction is focusing on the study of new modalities to get and to use the information deriving from our body in order to develop interfaces able to interact with computers. Currently there are various technologies for capturing the movement of the body. Today accelerometers are the most frequently devices used to do it, thanks to their availability on the market and to their modest price, but they are still far from being optimal. Actually accelerometers are able to capture only a small part of information posture and motion. Among other modalities for capturing movements, there is the *surface electromyography* (sEMG). This is a technique that derives from the medical field and is able to recognize the electrical activity produced by skeletal muscles. It is called “surface” because in the normal EMG investigation the muscles of the patient are reached by needles inserted directly inside the muscle, while in the sEMG the electrodes pad are used in place of needles. This paper presents a prototype of an sEMG interaction system. We have built a three-channel wearable electromyography. Furthermore we have built specific software that is able to acquire the digital data from the board, classify the movement and train an ANN and store the user's parameters and data in a MySQL DB in order to study how different parameters can

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 391–398, 2011.

affect the sEMG signal. In the first part of this paper we introduce how our body generates this specific electric signal and which parts of the body are involved. Afterwards we describe the characteristics of the sEMG information and the difficulties that this kind of technique involves. Then we describe how we have elaborated the signal, the extraction of the features and the classification of the movement, by using two different methods: wavelet analysis and independent component analysis. Next we describe the characteristic of the hardware and software system, and the setup for the system performance evaluation. Finally we show the results and present future developments of the wearable sEMG system.

2 Related Works

The new way to interact with computers or machines in general is by gestures and bodily motion; this originates from the face or hand. Many approaches have been made using cameras and computer vision algorithms to interpret sign languages. The most famous application in gesture recognition is probably the *Wiimote*. Furthermore there are systems based on one or two cameras that recognize a part of the body (typically the hands and the fingers), in which gestures are encoded as command for the computer, e.g. the S*ixthSense* from MIT Media Lab [7] that uses a camera and a tiny projector in a pendant-like device to see what is requested, and visually augment the surfaces or objects with which the user interacts by hand gesture, thanks to colored markers placed on the four fingers. Electromyography is a well-known technique in the medical field; it is used to diagnose pathology such as amyotrophic lateral sclerosis; otherwise it is used to facilitate amputees to command robot prosthesis [3]. Lately the techniques have been improved and now the control is made using not only one single muscle control, but using an entire group of muscles [7]. EMG signals can be used for a variety of applications including clinical applications, such as HCI and interactive computer gaming [4]. Other studies in the domain of bioengineering have concentrated on the use of electromyographic signals for controlling of prosthesis, rehabilitation and computer interfaces for users with motor disabilities [6] [10]. For example Benko and Saponas [1] presented a touch sensing developed through an interactive surface, and muscle sensing via EMG. Another work has been developed by Microsoft Research and Saponas et. al. [8] [9] in which is presented a muscle-sensing interface for always-available input in real-world applications (e.g. to interact with portable music player, videogames, car) able to classify and use in real-time a variety of finger gesture in order to interpret forearm electromyography (EMG) and classify finger gestures on a physical surface. They studied the existing taxonomies of natural human grips to develop a gesture set covering interaction in free space even when hands are busy with other objects.

3 The Electromyography System

The system consists of 3-channel wearable sEMG device that the user wears on his forearm by means of a forearm cover. The device is connected to an external device that provides the signal elaboration and allows the interaction of the user. We have built a prototype of a sEMG developing board, named *Eracle* (see Fig. 1) able to handle 3 different and independent channels (three triplets of electrodes).

It is equipped with PIC16F688 processor. In the first step the signals pass through the INA amplifier (the gain is adjustable from 200dB to 2000dB) and through a high-pass filter (cutoff frequency: 1.5Hz); afterwards an anti-aliasing filter is inserted, which is the Sallen-Key (double pole at 150Hz). This architecture is replicated for each EMG-channel. Then, the analog signals are digitalized from the microcontroller, which has a 10bit ADC (sampling rate: 270 Samples/Second); the samples are transferred to the computer via USB. The Eracle's output is composed of a string that contains for each row the data for each channel.

3.1 Data Elaboration

An external desktop computer performs the elaboration of the signal and the storage of the information that have been acquired. A standard desktop computer runs the software that elaborates the digital signals. It performs the training of the artificial neural network and hosts the MySQL DB, where the information is gathered.

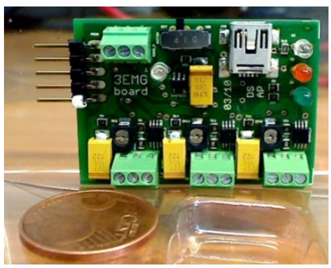


Fig. 1. The Wearable Eracle Acquisition Board compared with 5 cent coin

4 Signal Elaboration

This section presents how the electromyography signal is generated and the strategy developed for movement splitting, the methods for feature extraction and the way for movement classification (see Fig. 2). For elaborating the signals we have used two methods: the wavelet and independent component analysis (FastICA algorithm). The feature extractions have been performed with the same acquisition for both methods. EMG measures the electrical currents that are generated in a muscle during its contraction and represent neuromuscular activities. The neuromuscular system must be considered as an association of several functional units, called motor units (MUs). Nervous and muscular cells (muscle fibers) are excitable cells and their measured potential is stable within -70mV and -90mV. EMG signal detection is a delicate matter, and presents two main issues that strongly influence the fidelity of the signal. The first is the signal to noise ratio, defined as the ration between the energy of the EMG signal to the ratio of noise signal; and the distortion of the signal itself.

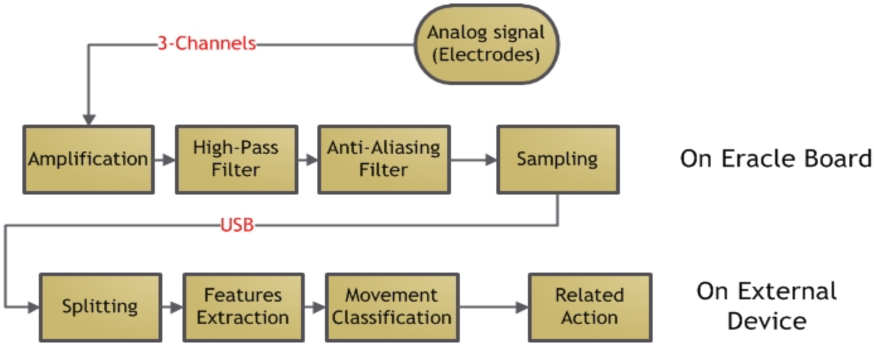


Fig. 2. DataStream of the application. The signal comes from the electrode pads; it passes through the acquisition board and then is acquired from the desktop station.

EMG signal is stochastic (random) in nature, and that can be represented by a Gaussian distribution function. The amplitude of the signal ranges from 0 to 10mV (peak-to-peak) or 0 to 1,5mV (rms). The usable energy of the signal is generally limited to a specific frequency range (0 to 500Hz), with the frequencies that most suits our purposes centered between 50 to 150Hz. Noise, instead, in the electromyographic techniques, may emanate from a wide range of sources: from the ambient (external disturbances), from our body (cross talking, endogenous disturbances) or from movement artifacts are the general noise related problems, which have to be addressed when working with EMG signal.

4.1 Wavelet Analysis

The wavelet is one of the best known of tool for signal analysis. The first step to perform for understanding the EMG signals is the splitting. The digital signal received from the EMG board is divided in many bursts, each related to a specific contraction in order to understand when a contraction starts and ends. An EMG signal is always made of a deterministic part, which contains the movement, and a part which contains the signal in its baseline configuration, when the muscles are inactive but the electrodes register a sort of basal activity and capture noise from environment. A neat separation using the raw signal is impossible, due to its high frequency oscillation that prevents it from using a fixed threshold to recognize when a movement is being executed. Working on the smoothed signal allows us to easily recognize the movement when it is performed, and so fixing a threshold on the elaborated signal leads to more effective results. This phase consists of three steps:

1. Rectification
2. Linear envelope
3. Segmentation

The feature extraction module has the role of identifying particular numeric parameters from the single signal burst. The integral EMG (linear envelope), moving absolute mean of the whole signal and its skewness has been used as feature as well. The Continuous Wavelet Transform (CWT) is here used. According to empiric observations and results obtained in other works, we have used for our analysis the Morlet mother wavelet [3]. Computing the CWT of a single burst composed by 270 samples produces a matrix of 5X270 size, and using the Singular Value Decomposition (SVD) we extract a vector of features that has seven components.

4.2 FastICA Analysis

In biometrics analysis the ICA methods are mainly employed for denoising. The main problem in using ICA methods with noisy signals is the lack of robustness of the algorithms; however, for sEMG data analysis, these approaches have been used with good results by Naik et al. [13]. From this point of view, ICA is also used to reduce the cross-talking between two or more muscles that are activated in the same contraction, thus the obtained independent components are the real signals generated by each motor unit. First of all the amplitude values of the three signals have been normalized. Moreover the signal sampled on channel 2 that is very noisy, maybe due to crosstalking has been limited by the ICA algorithm. On the contrary, the source signal sampled by channel 1 has been boosted, while the input signal on channel 3 remained virtually unchanged. Removing the noise from the signals it is possible to correctly extract some features, such as the root mean square (RMS), that identify the gesture performed by the user.

5 Software Module

The software module that has been developed for acquiring the electromyographic signal is able to acquire the digital data from board, classify the movements, train an ANN and store the user's parameters and data in MySQL DB in order to study how different parameters can affect the sEMG signal (see Fig. 3). First of all the user inserts his/her parameters in the form. The parameters stored so far are the following: Age, Gender, Height, Weight, Ethnic Group, Body Mass Index, Physical Activity, and Muscular Pathology. After the user data entry, the acquisition can be activated. For each movement, a short movie is presented, which shows how the user must perform the movement in order to improve the quality of the classification, and to avoid as much as possible errors in the signal splitting. After each acquisition the data are stored in a text file, and are automatically divided in a single file for each channel. Then it is possible to see the raw data acquired for each channel (the *"View Ch"* buttons). This is a first way of verifying the quality of the acquisition. Subsequently, it is possible to see the signal after the pre-elaboration (and to compare it with the raw signal), and before the launch of the classification and recognition methods. This elaboration consists in the rectification, the Butterworth filtering (in order to remove the main noise components [14]), the linear envelope of the filtered signal and then the splitting of the signal for each movement. These features allow us to check, straight afterwards every acquisition, if the quality of the data acquired is good. In case the quality is not good, it is possible to repeat the specific movement acquisition. After this we run a MATLAB program that performs the extraction of the features by using wavelet and ICA analysis, the classification of the neural network and the training session in order to test the quality of the whole session. Finally, the user's parameters and data are stored in a MySQL DB, so that it is possible to study how the different parameters can affect the sEMG signal. This is useful during the prototyping phase, because we can identify which are the common features of the target user for a specific application, or if the number of electrode pads are appropriate, or if their position inside the clothes is correct.

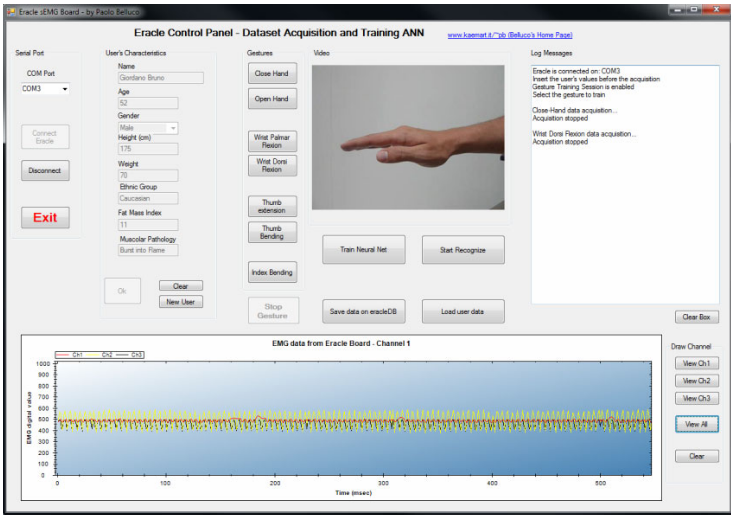


Fig. 2. The Eracle software Graphical User Interface

6 Evaluation Tests

We have organized some evaluation tests of the system. The time dedicated to each test is around 25 minutes, which include the placement of the electrodes, the storing of the users' parameters, the acquisition of the gestures, the elaboration by using the two methods and training of the ANN and finally the test for recognizing the gestures. The movements have been selected among those gestures that are considered easy to perform for a generic user. Particular attention has been devoted to verify that the selected movement performs well with the technology used, the techniques used for signal analysis adopted, and the kind of feature that has been extracted. In order to find a balance between the need of using 'user-friendly' movements and 'easy-to-analyze' movements, several tests have been performed on the same subject by evaluating different movements and estimating both the difficulties and the performances of the ANN in terms of number of movements correctly recognized. Experimentation confirms that the ANN with BP architecture for the classification, used in combination with the wavelets or FastICA, gives good performance for seven different movements, which are:

- open hand;
- close hand;
- wrist dorsi flexion;

- wrist palm flexion;
- thumb extension (only by wavelet);
- thumb bending (only by wavelet);
- index extension (together with the medium finger);

Table 1 shows that the movement recognition rate of a trained net using wavelet is very high, with a mean for all movements over 95%. By using FastICA the results are worse, but the sEMG recording classification is done almost in real-time.

Table 1. Success in percentage of movements recognized, referred to four subjects, computed by our system based on wavelet and FastICA

Movements	Wavelet %	FastICA %
Open	99.1	85
Close	95.6	90.2
Wrist Dorsi Flexion	97.8	92.1
Wrist Palm Flexion	99.3	89.5
Thumb Extension	95	-
Thumb Bending	98.4	-
Index	95.6	72.2

6.1 Discussion about the Results

The analyses conducted on the Eracle board reveal that the output digital signal, elaborated by the sEMG signals by using two different methods, is clean and suitable to our purposes despite the small dimensions of the board. The Eracle board is small enough to be inserted inside the clothes with minimum annoyance for the users. The performances related to the gesture recognition are good and the metaphors, implemented for the interaction, are powerful and intuitive. The movements recognized are easy to perform, and allow the user to act naturally even when he is using one of these movements as a command. Especially the movement related to thumb and index/medium has shown to be a good choice. In fact, the index is often used in the normal life to point at objects, and so it is suitable for commands in those user interfaces that include some kind of pointing action; and the thumb, the most important finger for human being, is often free to be used as a source of command.

7 Conclusion

This paper has presented a novel wearable interaction system based on the electric activity of the human muscles. The system, named Eracle, is a wearable acquisition system based on sEMG signals that allow us to recognize and manage the user's gestures by using two different promising methods (wavelet and FastICA) for features extraction. The effectiveness and performances of the system have been demonstrated through some users' tests. In the future, we will replace the electrodes with another type, which will sewed up inside the clothes, and we will also integrate other different signal acquiring techniques, based on the use of other kinds of sensors, in order to increase the number and the quality of the recognized gestures.

References

1. Benko, H., Saponas, T.S., Morris, D., Tan, D.: Enhancing input on and above the interactive surface with muscle sensing. In: ITS 2009, Banff, Alberta, Canada (2009)
2. Goupillaud, P., Grossmann, A., Morlet, J.: Cycle-octave and related transforms in seismic signal analysis. Geoexploration, Seismic Signal Analysis and Discrimination III 23(1), 85–102 (1984)
3. Hogan, N., Mann, R.W.: Myoelectric signal processing: optimal estimation applied to electromyography part I: derivation of the optimal myoprocessor. IEEE Transactions on Biomedical Engineering 27(7), 382–395 (1980)
4. Kim, J., Bee, N., Wagner, J., André, E.: Emote to win: Affective interactions with a computer game agent. Lecture Notes in Informatics (LNI), vol. 50, pp. 159–164 (September 2004)
5. Maes, P.: Sixthsense: Integrating information and the real world. In: Proc. 8th IEEE Int. Symp. on Mixed and Augmented Reality - ISMAR 2009 (2009)
6. Moon, I., Lee, M., Ryu, J., Mun, M.: Intelligent robotic wheelchair with emg-, gesture-, and voice-based interfaces. In: Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS 2003), vol. 4, pp. 3453–3458 (2003)
7. Parker, P., Englehart, K., Hudgins, B.: Myoelectric signal processing for control of powered limb prostheses. J. Electromyogr. Kinesiol. 16(6), 541–548 (2006)
8. Saponas, T.S., Tan, D.S., Morris, D., Balakrishnan, R.: Demonstrating the feasibility of using forearm electromyography for muscle-computer interfaces. In: Proc. CHI 2008, New York, NY, USA, pp. 515–524 (2008)
9. Saponas, T.S., Tan, D.S., Morris, D., Balakrishnan, R., Turner, J., Landay, J.A.: Enabling always-available input with muscle-computer interfaces. In: Proc. UIST 2009, New York, NY, USA, pp. 167–176 (2009)
10. Yamaguchi, H., Nishikawa, D., Yu, W., Maruishi, M., Yokoi, H., Mano, Y., Kakazu, Y.: Emg control switching system for fes. In: Proc. 8th IEEE Int. Workshop Robot and Human Interaction - RO-MAN 1999, pp. 7–12 (1999)

Development of Tactile and Haptic Systems for U.S. Infantry Navigation and Communication

Linda R. Elliott[1], Elmar T. Schmeisser[2], and Elizabeth S. Redden[1]

[1] US Army Research Laboratory, Human Research and Engineering Directorate, Maneuver Center of Excellence, Fort Benning, GA, USA
[2] US Army Research Office, LS/Neurosciences, Research Triangle Park, NC, USA
linda.r.elliott@us.army.mil

Abstract. In this paper we discuss plans initiated to develop and evaluate multisensory displays (i.e. visual, haptic, tactile) to support dismounted (i.e., not in vehicle) Soldier movement, communication, and targeting. Human factors studies of an array of military operational roles have shown significant demand for focal visual attention that diminishes the capacity for task-sharing and attention allocation, especially in the context of unexpected changes and events. If other sensory modalities can be effectively used in a military environment, the benefit could be significant in increasing survivability, information flow, and mission achievement. We discuss operational task demands and two efforts supported from a 2010 SBIR (Small Business Innovative Research) topic.

Keywords: Tactile displays; Haptic displays; Soldier navigation; Soldier performance; Multisensory displays; Intuitive displays.

1 Introduction

Tactile and haptic interfaces have long been used for intuitive displays and controls. Consider our daily use of devices such as steering wheels, joysticks, and computer mice (haptic), and cell phone vibrating alerts (tactile). Haptic devices incorporate the sense of touch and/or kinesthesis from motor activity based in the skin, muscles, joints and tendons, and includes tactile sensing [1]. Tactile devices are thus haptic by definition, but are based more specifically on stimulation of the skin—usually mechanical, but also including chemical or heat stimulations. It should be noted that haptic, and sometimes tactile, interactions are often in the form of input-output feedback loops, as opposed to more passive reception. That is, the user is actively interacting with the device.

While simple versions of haptic and tactile devices are commonplace, more sophisticated applications are also proving their worth, in areas such as virtual reality [2], [3], robotic telepresence [4], and spatial orientation [5]. At the same time, research has focused on a number of human factors issues regarding multisensory display, such as sensory adaptation, spatial masking, temporal masking, and cross-modal interactions [6]. A series of meta-analyses comprising hundreds of experiment-based comparisons, has shown significant advantages can be gained from addition of

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 399–407, 2011.

tactile information to existing visual displays [7]. In this paper, we will describe some applications of haptic and tactile devices to military performance, particularly Soldier performance, and plans for further instantiations of this capability as it can apply to combat Soldiers.

1.1 Military Applications

Haptic Systems. Haptic systems for military applications include devices such as haptic joysticks and gloves for control of robots and exoskeletons [8]. Haptic-controlled exoskeletons have been demonstrated to ease physical workload and also to aid in rehabilitation of patients with stroke or head injuries [9]. Haptic capabilities have been demonstrated in virtual reality simulations that can be used for training or telepresence [10]. In addition, gesture-based controls can also be considered haptic in nature.

Tactile Systems. Tactile systems for military performance have demonstrated their potential with regard to capability achievement and performance advantage, across a number of applications. Experiments and demonstrations have been conducted across a wide range of settings, from laboratory tasks to high-fidelity simulations and real-world environments. Operators of these various tactile systems have successfully perceived and interpreted vibrotactile cues in averse, demanding, and distracting situations, such as combat vehicles [11], aircrew cockpits [12], high-speed watercraft [13], underwater environments [14] and during strenuous movements [15]. Also, tactile alerts have effectively supported performance during robot control [16] and UAV control [17].

1.2 Some ARL Soldier-Based Experiments of Tactile Displays

While tactile cues have been associated with better or faster performance, we cannot assume that all tactile cue displays will be effective. However, if we consider predictions of performance gained through (a) alleviation of sensory overload [18] or (b) alleviation of cognitive deliberation [6] we should expect that given a high workload multitask and multisensory situation, that a multisensory approach including a tactile channel will reduce workload and improve performance, assuming the tactile display is easily perceived and comprehended.

Several HRED experiments conducted with Soldiers support these expectations. We started with task analysis information that identified key situations in which Soldiers are visually overloaded [19] then initiated several HRED studies to investigate effects of tactile cues in context [20]. The studies further supported the ability of Soldiers to detect not only single alerts but also patterns of multiple tactors to represent different communications [15]. It is particularly promising that these patterns can be perceived during strenuous movements [21].

Because tactile alerts have been demonstrated to effectively and easily portray spatial orientation [5] it is reasonable to assume that a torso belt would convey direction information that is also immediately understood. Three experiments investigated the efficacy and suitability of a torso-mounted tactile belt for Soldier navigation [22]. For all experiments, researchers used a tactile navigation system developed by the Netherlands Organisation for Applied Scientific Research – TNO

(Toegepast-Natuurwetenschappelijk Onderzoek), for U.S. Army Soldier land navigation (see Figure 1). Land navigation occurred at the Fort Benning training site for Soldier navigation – a challenging terrain that requires careful attention to one's surroundings for threat from terrain and natural wildlife. In experiment 1 participants navigated three waypoints along 600 meters through heavily wooded terrain, using each system: the TNO tactile system, a map and compass, and with a handheld Army GPS system. Performance measures included quantitative measurement of navigation error, along with other objective and Soldier-based assessments (see figure 2).

Fig. 1. Personal Tactile Navigator System

In a second experiment researchers used terrain similar to the first experiment with the additional challenge of night operations during intermittent thunderstorms. Soldiers navigated three waypoints while also searching for live and silhouette targets, using (a) handheld GPS device, (b) head-mounted map-based GPS, and (c) the tactile GPS system. Experiment 3 had participants navigate with (a) a commercial GPS arrow display, (b) the tactile GPS system, and (c) both together. Given this series of results, more extensively discussed elsewhere [22], it was concluded that tactile navigation displays can be used in strenuous outdoor environments and can outperform visual displays under conditions of high cognitive and visual workload. In addition, Soldier comments were highly in favor of this capability, stating that the system was “hands-free, eyes-free, and mind-free.”

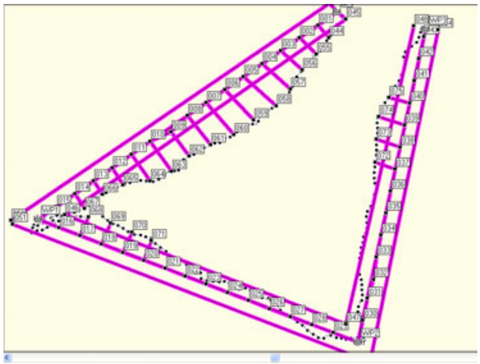


Fig. 2. Quantitative MapQuest-based approach to performance measurement (navigation)

A torso-mounted tactile belt was also demonstrated as useful for robot control operations [23]. In this experiment, three types of robot controller navigation map display configurations were evaluated for effects on beyond line-of-sight robotic navigation tasks. First was a larger split screen visual display that presented both a map display and a camera-based driving display on a 6.5 inch screen. Two smaller alternatives were also evaluated. One alternative was a 3.5 inch display that allowed the operator to toggle back and forth between the driving display and the map display. The third option added a torso-mounted tactile display to the toggle-based display in order to provide direction information simultaneously with the camera display and thus reduce the need to toggle as frequently to the map display. Each display option was evaluated based on objective performance data, expert-based observations, and scaled subjective Soldier questionnaire items. Findings indicated that operators' navigation performance with the smaller, toggle display was much worse than with the larger split screen display. However, when the tactile display was added to the toggle display, performance was as effective as with the larger display.

2 SBIR Topic

The experiments described above establish the potential of tactile systems for supporting Soldier performance while easing workload and gaining high user acceptance. At the same time, Soldiers provided many suggestions for device design before a system can be practically used in combat. Certainly, the device must be made to be lightweight, comfortable, rugged, and easily maintained. The device must enable wireless communication among Soldiers, to enable commanders to easily and covertly signal Soldiers regarding alerts or movements. This would build upon battlefield visualization techniques now common to command and control, by enabling the commanders to quickly relate critical communications as to where to go or where to shoot. In this way, the integration of a visual command center with distributed tactile communications enable dynamic battle maneuvers with intuitively understood signals.

At the same time, we also need the development of tactile systems that can be used as research testbeds, to enable further research in multisensory performance issues while being fully grounded to Soldier task demands. Thus, systems must also provide the means by which performance can be easily assessed—tracking communications, time-stamped performance events, GPS-enabled assessment of navigation performance, user-based assessments of situation awareness, and data logs. The purpose of the SBIR topic was to solicit approaches to such a tactile system—one that can serve to illustrate the advantage to Soldier performance, while offering a testbed for research. Of course, the resulting capabilities would generalize to many other military, government, first-responder, and commercial (e.g., hiking, hunting, tourist navigation) applications.

2.1 Tactile System with Navigation guidance

Engineering Acoustics, Inc. (EAI) has a long history of tactile system engineering, for many military applications—such as situation awareness support for aircraft pilots [24]. Guided by experiments, sensory perception research, and user requirements, EAI has engineered specialized tactors that are very easily perceived by users, more accurately than other types of tactors, even during combat movements [21]. The military environment is severe, placing demands on equipment, system and Soldier. War-fighters may be under conditions of extreme physical noise and stress and, in critical situations, cannot afford to "miss" information. As part of this SBIR effort, EAI will demonstrate advanced tactors designed to overcome problems that can occur (e.g., poor perception, slow spin-up times, vulnerability to loading). Particularly, they will attend to the problem of auditory signatures, to enable a system that is more completely covert. The potential of a wearable vibrotactile system has been impeded by a lack of suitable tactors and sensors and non standardized tactile symbology. Figure 3 shows the EAI tactor types and steady state displacement performance into a skin load.

Figure 4 shows a block diagram for the system. The system comprises visual display hardware(for Phase I we propose using a handheld or Tablet computer for this task) , EAI tactor controller and belt array, a COTS (Commercial off the shelf) GPS / compass sensor interfaced directly to the tactor controller and software components.

In addition to engineering the system, EAI will address human factors considerations, building upon accumulated knowledge, with particular focus on Soldier requirements. The use of multiple tactile locations in vibrotactile displays for spatial orientation is known to be intuitive; however, effective design must, from the start, be based on user task demands and the context of performance. While the system will be constructed toward user requirements, it will also enable systematic research with regard to impact of multisensory systems on workload, grounded towards the context of Soldier navigation, communication, and performance. Thus the system and the research it enables are expected to support the core performance domain for Soldiers – to move, to shoot, and to communicate.

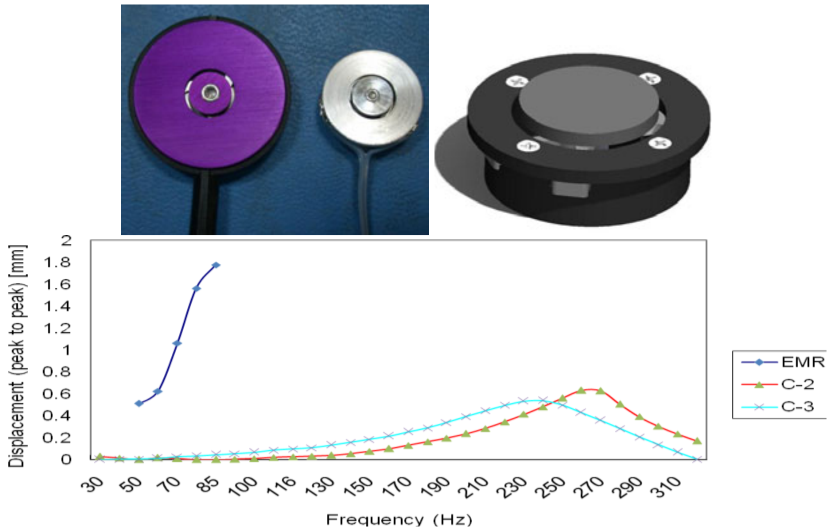


Fig. 3. The EAI C-2, C-3 and EMR vibrotactor transducers (top left to right respectively) together with their operational displacement output (measured with an optical fiber displacement sensor on a silicone skin phantom)

Fig. 4. Block diagram for the proposed ATAC-NavCom system

2.2 Tactile System with Haptic Communications

The COMMAND (Communication-based Operational Multi-Modal Automated Navigation Device) system, in development by Anthrotronix, will build upon their previous work for Office of Naval Research, which led to the development of a Haptic Automated Communication System (HACS) that utilizes an instrumented glove for real-time communications based on hand signals. The instrumented glove includes 6 embedded accelerometers, a gyroscope, and a digital compass for automated recognition of standard hand and arm signals, gestures, pointing, and weapon firing. It also includes a torso-mounted accelerometer and digital compass for Soldier location and stance (e.g., upright, prone). Also included is a haptic display vest with 20 tactors for pattern-based communications, and a GPS-enabled handheld computer (see Fig 5).

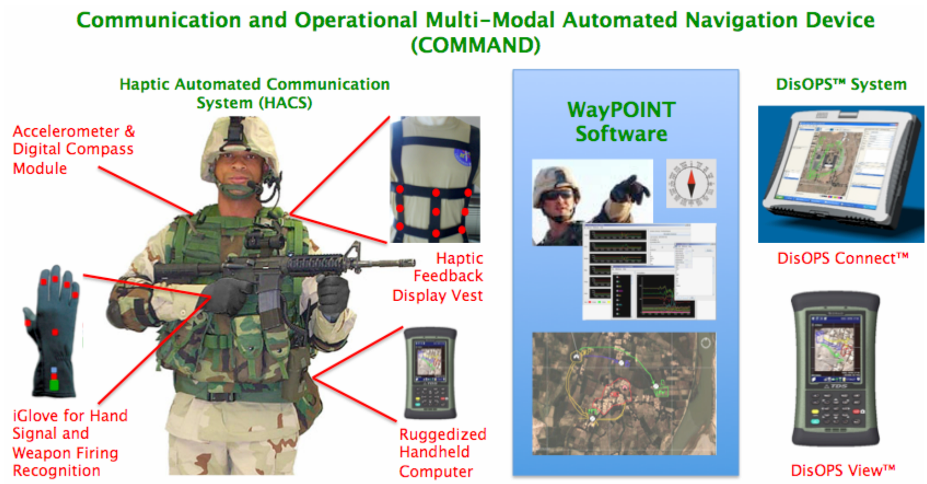


Fig. 5. COMMAND System

While initially the system will demonstrate a proof of concept, there are plans to eventually provide real-time geospatial reports and tools for pre-mission planning and review. This would entail integration with the Lockheed Martin DisOPS™ (Distributed Operations) System, which is composed of a software package and a ruggedized handheld computing system, visualized to be used by Army squad and fire team leaders.

The COMMAND system is visualized to integrate these technologies and incorporate a novel software system (WayPOINT) to provide intelligent, automated, multimodal information processing and display to support dismount infantry operations. The resulting system will enable gesture-based communications, as well as tactile-based navigation and communications.

3 Discussion

Multiple experiments and demonstrations have proven the theory-based predictions regarding advantages of haptic and tactile cues to support performance in high-workload situations, particularly multi-tasked situations with high demands for focal visual attention. Task analysis models identified that Soldiers have very high demands for visual attention, particularly when Soldiers are moving or shooting. Subsequent experiments proved the value of tactile systems to support Soldier navigation and communication. At the same time, systems must be improved and refined before they can be practical in combat situations. They must be made lightweight, comfortable, rugged, networked within a command and control system, and they must be easy to use and easy to maintain. As tactile displays are increasingly used for communication of more complex and multiple concepts, it will become evident that tactile and multisensory systems in general must be designed for rapid and easy comprehension. This paper described efforts underway toward the goal of effective support of Soldier performance, and the development of a system that can also be used for grounded research (i.e.,high generalizability to military operations) in multisensory perception and comprehension.

References

1. Van Erp, J.B.F., Kyung, K., Kassner, S., Carter, J., Brewster, S., Weber, G., Andrew, A.: Setting the Standards for Haptic and Tactile Interactions: ISO's Work. In: Proceedings of Eurohaptics 2010 Conference, Amsterdam, The Netherlands (2010)
2. Iwata, H.: Haptic Interfaces. In: Jacko, J.A., Sears, A. (eds.) The Human-Computer Interaction Handbook, pp. 207–219. Lawrence Erlbaum Associates, London (2003)
3. Van Erp, J.B.: Tactile displays in virtual environments. Report No. ADP010788. Soesterberg, The Netherlands: TNO Human Factors Research Institute (2001)
4. Van Erp, J., Duistermaat, M., Jansen, C., Groen, E., Hoedemaeker, M.: Telepresence: Bringing the Operator back in the loop. RTO-MP-HFM-136 (2006)
5. Chiasson, J., McGrath, B., Rupert, A.: Enhanced situation awareness in sea, air, and land environment. In: Proceedings of NATO RTO Human Factors & Medicine Panel Symposium on "Spatial Disorientation in Military Vehicles: Causes, Consequences and Cures, La Coruña, Spain, No. TRO-MP-086, pp. 1–10 (2002)
6. Van Erp, J.: Tactile displays for navigation and orientation: Perception and behavior. Mostert & Van Onderen, Leiden (2007)
7. Elliott, L., Coovert, M., Redden, E.: A summary review of meta-analysis of tactile and visual displays. Invited paper Presented to the 13th International Conference on Human-Computer Interaction, San Diego, CA (June 2009)
8. Greenemeier, L.: The future of exoskeletons: Lighter loads, limbs, and more. Scientific American (2007), http://www.scientificamerican.com/article.cfm?id=the-future-of-exoskeleton (February 4, 2011)
9. Rocon, E., Pons, J.: Exoskeletons in rehabilitation robotics. Springer Tracts in Advanced Robotics, vol. 69. Springer, Heidelberg (2011)
10. Aliberti, D., Bruen, T.L.: Telepresence: Harnessing the human-computer-machine interface. Army Logistician 38(1) (2006)
11. Carlander, O., Errikson, L.: Uni- and bimodal threat cueing with vibrotactile and 3-D audio technologies in a combat vehicle. In: Proceedings of the Human Factors and Ergonomics 50th Annual Meeting, pp. 1552–1556. Human Factors and Ergonomic Society, Santa Monica (2006)

12. Rupert, A.H., Graithwaite, M., McGrath, B., Estrada, A., Raj, A.: Tactile Situation Awareness System Flight Demonstration (2004), http://www.stormingmedia.us/corpauthors/ARMY_AEROMEDICAL_RESEARCH_LAB_FORT_RUCKER_AL.html Army Aeromedical Resaerch Lab, Fort Rucker, AL. Report No. A891224
13. Dobbins, T., Samways, S.: The use of tactile navigation cues in high-speed craft operations. In: Proceedings of the RINA Conference on High Speed Craft: Technology and Operation, pp. 13–20. The royal Institution of Naval Architects, London (2002)
14. Self, B., van Erp, J., Eriksson, L., Elliott, L.R.: Human factors issues of tactile displays for military environments. In: J. van Erp (Ed.) Tactile Displays for Military Environments. NATO (2007) Report No. RTO-TR-HFM-122
15. Pettitt, R., Redden, E.S., Carstens, C.: Comparison of army hand and arm signals to a covert tactile communication system in a dynamic environment. Report No. ARL-TR-3838. Aberdeen Proving Ground, Army Research Laboratory, MD (2006)
16. Redden, E., Elliott, L., Pettitt, R., Carstens, C.: A tactile option to reduce robot controller size. Journal on Multimodal User Interfaces 2(34), 205–216 (2009)
17. McKinley, R., Albery, W., Tripp, L.: Multisensory cueing to improve UAV operator performance during landing. In: Proceedings of the NATO HFM Symposium, Human Factors of Uninhabited Military Vehicles as Force Multipliers, Biaritz, France. RTO-MP-HFM-135, paper 20. NATO RTO, Neuilly-sur-Seine Cedex (2006)
18. Wickens, C.: Multiple resources and mental workload. Human Factors 50(3), 449–454 (2008)
19. Mitchell, D., Samms, C., Glumm, M., Krausman, A., Brelsford, M., Garrett, L.: Improved Performance Research Integration Tool (IMPRINT) Model Analyses in Support of the Situational Understanding as an Enabler for Unit of Action Maneuver Team Soldiers Science and Technology Objective (STO) in support of Future Combat Systems (FCS). Aberdeen Proving Ground, MD: US Army Research Laboratory (2004)
20. Elliott, L.R., Redden, E., Krausman, A., Carstens, C.: Multi-modal displays to support Army Infantry Decisionmaking and performance. In: Proceedings of the 2005 International Conference on Naturalistic Decisionmaking, Amsterdam, NL (June 2005)
21. Redden, E.S., Carstens, C.B., Turner, D.D., Elliott, L.R.: Localization of Tactile Signals as a Function of Tactor Operating Characteristics (Technical Report ARL-TR-3971.) Aberdeen Proving Ground, MD: US Army Research Laboratory (2006)
22. Elliott, L., van Erp, J., Redden, E., Duistermaat, M.: Field-Based Validation of a Tactile Navigation Device. IEEE Transactions on Haptics 3(2), 78–87 (2010)
23. Rupert, A.H., Guedry, F.E., Reschke, M.F.: The use of a tactile interface to convey position and motion perceptions. AGARD Meeting on Virtual Interfaces: Research and Applications (October 1993)

Utilization of Shadow Media - Supporting Co-Creation of Bodily Expression Activity in a Group

Koji Iida[1], Shiroh Itai[1], Hiroko Nishi[2], and Yoshiyuki Miwa[1]

[1] Waseda University, 3-4-1, Ohkubo, Shinjuku-ku, Tokyo, Japan
[2] Toyo Eiwa University, 32, Miho-cho, Midori-ku, Yokohama, Kanagawa, Japan
koji.e-da@suou.waseda.jp, itai@fuji.waseda.jp,
hiroko@toyoeiwa.ac.jp, miwa@waseda.jp

Abstract. In this research, in order to assist creation of the body expression acidity in a group, the shadow media system that can cope with a group of up to 30 kindergarteners was developed and utilized in the kindergarten. As a result, we found out that the shadow media system can open up bodies of individuals to improve physical sensitivities or that body expression to be created differs depending on each shadow media type. Moreover, it was observed that how the individual body expression was evolved to the body expression activities in one coherent group. Additionally, it was revealed that formation or breakup of a group can be observed in the body expression activity in the group of such kindergarteners.

Keywords: Bodily Expression, Image, Awareness, Co-creation, Shadow Media.

1 Introduction

The authors have developed the shadow media system which supports creation of bodily expression by transforming shadows, which is inseparable from the body, into various shapes, giving a gap between the body and the shadow so as to encourage awareness [1]. Actually, when we transform the shadow of the body into string or polygon-shaped through image processing, we can observe motions and bodily sense which cannot be seen normally are drawn out unconsciously and the bodily expression is promoted [2]. Therefore, if we can utilize the shadow media which has such effects in the field of bodily expression activity in groups, individual creativity is enriched and a new world may be born by touching other's expressions and by co-creation of expressions. Especially in early age, expression means creating one's own world. It is an important factor which supports a small child when he creates a stage on which he can feel alive [3]. On the other hand, however, there is a growing concern that overflowing information media degrades capability of infant's bodily expression. Under such circumstances, as one of the methods which further enriches bodily expression, people are paying attention to technology which supports expression by utilizing visual and audio media which have been advancing rapidly these days in addition to conventional physical support [4-6]. However, this technology is

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 408–417, 2011.

only targeted at an individual or a few people. No research targeted at 20 to 30 people has been done so far. Therefore, there are few media technologies for expression which we can apply to the field of the bodily expression activity in which the whole class from kindergarten or elementary school will participate. In this regard, unlimited number of people principally can use this shadow media system. However, it had been hard to utilize the shadow media system to support bodily expression activities in a group of over 5 to 6 people because the size of projecting space is limited. Consequently, in this study, we decided to target educational scene of children and to develop a new version of shadow media system which can respond to maximum of 30 people. Additionally, once the developed system was applied to the scene, bodily sense had been improved because his or her body was opened by the shadow media, various shapes of the shadow media encouraged children create different images and bodily motions and connection between individual children was facilitated and a well-coordinated bodily expression was emerged. In short, the effectiveness of the shadow media was confirmed, which is reported as follows.

2 Development of Shadow Media System

The range of the shadow of human bodies projected in the shadow media system (size of the shadow media space) developed earlier shown in Fig. 1 is 4.0m in width, 3.0m in depth, and 3.0m in height. When we conduct the shadow media projection in the wide space (width 12m x depth 2.5m x height 2.0m) where 10 to 20 kindergartens can simultaneously experience, there will be following problems.

(a) The distance from the screen to the thermo camera that captures human images (hereinafter called “acquisition distance”) gets longer.

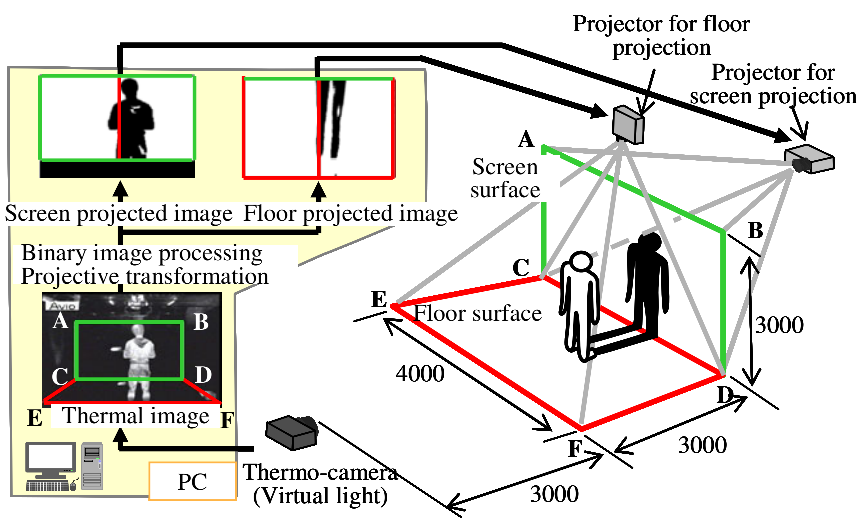


Fig. 1. Shadow media system

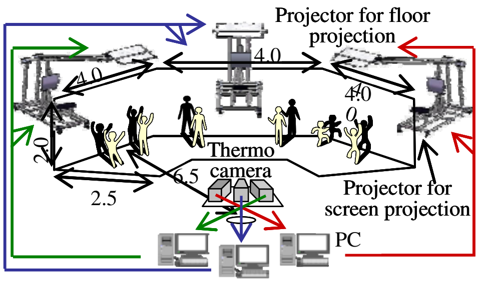


Fig. 2. Shadow medial for group activity

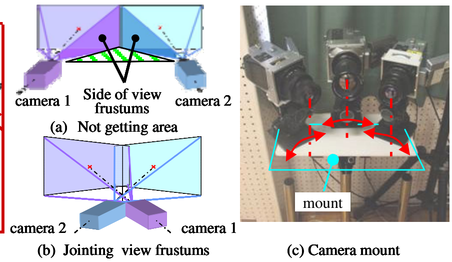


Fig. 3. Expanding angle of view

(b) The distance from the picture plane (floor or screen) to the projector (hereinafter called "projection distance") gets longer.

To solve these problems, we developed the device shown in Fig. 2 in this study. Firstly, for the solution of (a), we built the human image capturing device composed of multiple thermo cameras. It allows extension of the horizontal angle, thereby shortening the acquisition distance. Secondly, to solve problem (b), we developed the device resembling a mobile basketball goal as a shadow projection device. Through multiple combination of the device, range of the shadow projection is divided by several projectors, thus allowing shortening of the projection distance. We will explain the detail of these devices below.

Firstly, we will explain the device for capturing human images. If we deploy the thermo cameras as shown in Fig. 3(a), there will be an area (shaded area) where thermo cameras cannot capture human images. So, as shown in Fig. 3(b) we deployed the thermo cameras to overlap the view frustum of the neighboring thermo camera. In this device, we installed up to 3 thermo camera as shown in Fig. 3(c), and developed the camera mount that allows rotation of these cameras around the vertical axis passing through the optical center of the lens. This made it possible to expand the angle up to 160deg horizontally and 52deg verticality while connecting the image capturing range in seamless manner. As a result, we achieved capturing of the images of people in the space of 12m in width, 2.5m in depth, and 2.0m in height with the thermo camera installed around 6.5m from the screen.

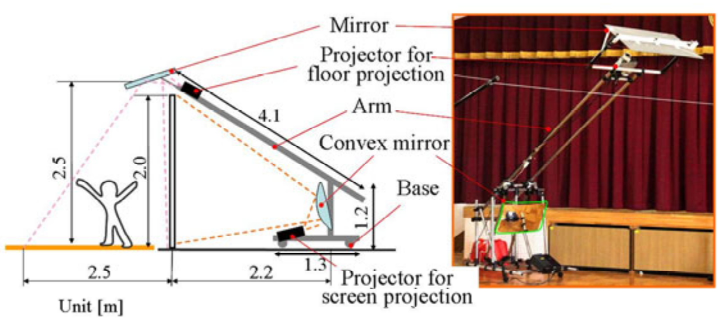


Fig. 4. Structure of projection device

Secondly, we will describe the shadow projection device resembling a mobile basketball goal. Fig. 4 shows the detail of this device. This device uses stainless steel pipes (38mm in diameter) as the constructional member. The device consists of two units: one is the truss structure basis (0.8m in width, 1.3m in depth, and 1.2m in height). The other is the arm (4.1m in length and 0.4m in width) extending obliquely upward on this basis. We can raise this arm though the counter weight by using single-degree rotating holder installed in the joint part of the two units as the fulcrum point. For the projection on the floor, the mirror (0.9m in width and 0.6m in height) installed on the distal end of the arm and the projector installed on the arm facing the mirror are used. In addition, the projector (BenQ MP771, horizontal angle of view: 73.9deg, vertical angle of view: 44.7deg) installed at the bottom of the base archived rear projection on the screen through the convex mirror, decreasing the projection distance to 65% of conventional distance while expanding the projection range. Thus, one projector for the floor projection and other projector for the screen projection made it possible to project the shadow in the space of up to 5.0m in width, 2.5m in depth, and 2.0m in height. Finally, we achieved continuous projection of shadows of the bodies in the space of 12m in width, 2.5m in depth, and 2.0m in height in a continuous manner.

3 Bodily Expression Activity in a Group through the Shadow Media

3.1 Experiment Method

As shown in Fig. 5, a shadow media system designed to function for a group activity, which is previously explained, was installed in the hall (12.5m width, 10.8m depth, 3.0m height) at Daini Sawarabi kindergarten (Okazaki, Japan). A maximum of 30 kindergarten children, who make a group living in the same classroom every day, experience this system. Each party experiences this system for about 30 minutes. During the experiment, kindergarten children show bodily expression activity through regular shadow, string-shaped shadow, polygon shadow, dual residual shadow and particle shadow (please see reference [1] for the information about the dual residual shadow and the particle shadow). When conducting this experiment, our facilitator explains to kindergarten children that their shadow, deformed occasionally, is projected by this system. And kindergarten children freely performed the activity in the remaining time. He sometimes tells them to pay attention to their shadow which the

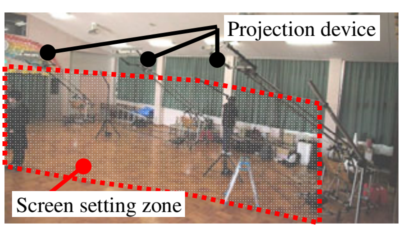


Fig. 5. Installation of the new shadow media system at kindergarten

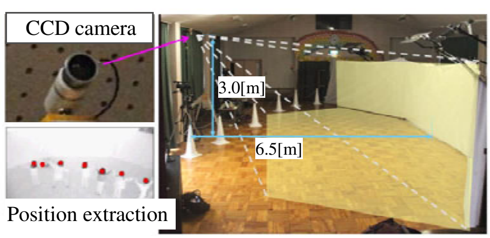


Fig. 6. Position measurement system

system projects in order to facilitate their bodily expression activity during this experiment.

We used video footage and simplified position-measurement method with CCD camera to carefully look into changes in individual bodily expression activity which is shown by the children. In the same manner, we also observed how bodily expression activity in a group of more than 5 children develops. For the measurement, as shown in Fig. 6, we installed a CCD camera, which had viewing angle of 135deg by 110deg and was equipped with fish-eye lens and infrared light transmission filter, at the level of 3.0m above the room floor to extract a human head region that does not reflect infrared light from an acquired image by digitalization processing. Furthermore, after the experiment of this system, we interviewed the children to hear what kind of things they expressed during the play. We also conducted an interview with their homeroom teachers.

3.2 Bodily Expression Activity by Individuals

First, a state of bodily expression activity drawn out by the shadow media in individuals is described. For the regular shadow which is not transformed, the way to use the shadows which change in sizes responding to the distance between the light source and the body was observed. When the regular shadow which had been artificially created was deleted, the kindergarten children became bewildered; they became either restless or motionless. It indicates that the children are highly interested in the shadows of their bodies.

Next, when the shadow was transformed, active state was voluntarily created without being encouraged by the facilitator and bodily expression which was not observed under the condition with regular shadow was seen. For example, when a string-shaped shadow was projected, there were many children who started wiggling their entire bodies like octopus and squid (Fig. 7(a)). When a polygon shadow was projected, some children widely stretched their arms and legs like star and starfish (Fig. 7(b)) while some children spread their arms up and down; other children imitating such motions. Furthermore, there was a scene where 2 to 3 children made one bodily expression together. For example, in Fig. 8, children made one shape together with the facilitator or moved together holding their hands while the s string-shaped shadow was projected. Also, there is a report on comments of homeroom teacher such as: “I was able to find a new aspect of a child who is normally quiet was actively engaged in bodily expression”, “it was surprising to see the children who usually do not play together worked together while experiencing this shadow media.” Note that the Table 1 shows a summary of what the children tried expressing obtained from the hearing investigation from them after the experience. From above results, we know that various images were evoked and multiple bodily expressions were drawn through the shadow media.

3.3 Bodily Expression Activity by Groups

Second, bodily expression activity in groups is described. Various bodily expression activities by 1 to 3 kindergarten children mentioned in previous section were further developed to activities of groups with more than 5 children. For example, as indicated

(a) String-shaped shadow (b) Polygon shadow

Fig. 7. Appearance of individual body expression

Fig. 8. Appearance of body expression in a few people

Table 1. Something that is expressed by infants

String-shaped shadow	Cactus, Rod-like person, Snake, Skeleton, Octopus, Squid, Eel, Shark
Polygon shadow	Star, Starfish, House, Heart-shape, Shark, Drill
Dual residual shadow	Riffle, Extinguishing fire, Sea, Invisible person, Fire, Magma, Bonfire
Particle shadow	Snow, Rainbow, Twinkling stars, Rain, Flower, Confetti

in Fig. 9, when string-shaped shadow is projected, each child wiggled their whole bodies discretely and expressed aquatic animals such as squid and octopus. Under such situation, one child spread his arms up and down, expressing a shark eating other aquatic animals with its wide mouth. Afterward, the boy took a lead and suggested other boys to express one large shark together: 4 boys moved across the space together while each of them representing the mouth, the body, and the tail respectively, creating one large shark swimming in the sea. Then, other children started joining the shark one after another, and eventually all the children experiencing the shadow media developed one gigantic shark swimming in the space. Later, when the gigantic shark became apart, some children continued to express shark while others trying to escape from the shark. Moreover, how a story of the body expression activity in a group was evolved was observed. Fig. 10 shows the children expressing a small fire lit with a match when dual residual shadow is projected. They first bent their bodies and then slowly started rising to express a spreading fire; intense frame was expressed by movement of each child making shape of sparks. Then the children display a motion to extinguish the fire with water in a bucket. That is to say, there was role allotment made naturally (i.e. some children were expressing fire while others expressing water) among the children in a series of expressions; they improvisationally made bodily expressions that matched the story. In conclusion, we found out that in the bodily expression activities of child groups using the shadow media, the image of each child was voluntarily shared and bodily expression of each was improvisationally developing into those of groups with co-creation of the story.

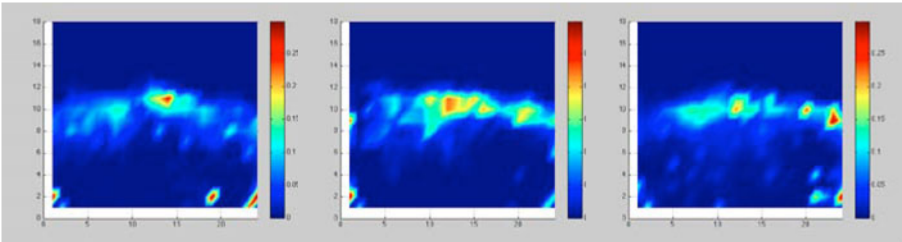
Fig. 9. Development into group activity from individual activity

Fig. 10. Story-making process in the group activity

3.4 Group Formation in Bodily Expression Activities

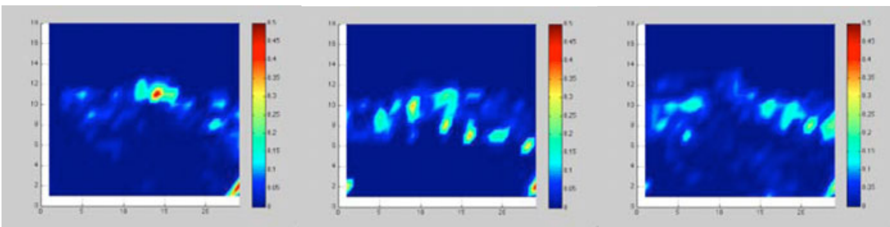
Next, result of position measurement of groups in bodily expression activities of the kindergarten children is described. To begin with, we have studied about the position of the shadow media space (width 12m x depth 2.5m) in which the frequency of bodily expression activities by the children was high. In order to do that, we have divided this space into grids, and observed the position of the grids each child stayed every 1/20 second. Note that the size of the grid is fixed in certain size (width 0.5m x depth 0.5m) to prevent several children from staying in the same grid. Then accumulated time of the children performed bodily expression in each grid was observed, and this accumulated time was calculated as the ratio against the presentation time of the shadow media. Presentation times of each shadow media are: 5 minutes for regular and polygon shadows, and 4 minutes for string-shaped shadow. Fig. 11 shows comparison of each shadow media indicating frequent distribution of the children in the shadow media space calculated by the aforementioned method. In this figure, red area indicates the area with high frequency rate of the children staying, and blue area indicates the area with low frequency rate of children staying. As a result, we found out that in any media, the children stayed 1m away from the screen, and moved widely from end to end of the shadow media space to perform bodily expression, in parallel to the screen.

Furthermore, we have studied group formation of the children in bodily expression activities. For this purpose, we decided to conduct calculation of frequent distribution of the children in abovementioned shadow media space every 30 seconds. It clarifies temporal transition of area with high frequency rate of the children staying as well as the number, allowing us to study temporal variation in group formation of the children. In regular shadow, the plurality of red regions where the children highly frequency stay are rarely found in any time zone, hence, we assume that the children uniformly spread in parallel to the screen and performed bodily expression (Fig. 12). On the other hand, in transformed shadows (string-shaped and polygon shadows), position and number of the red area were temporally changed as shown in Fig. 13 and 14. For example,



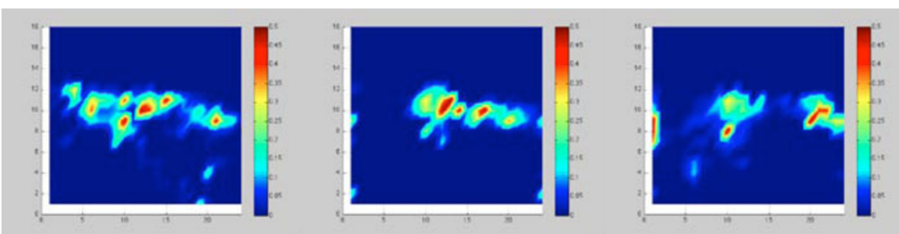
(a) Regular shadow (b) String-shaped shadow (c) Polygon shadow

Fig. 11. Frequent distribution of staying time in shadow media space



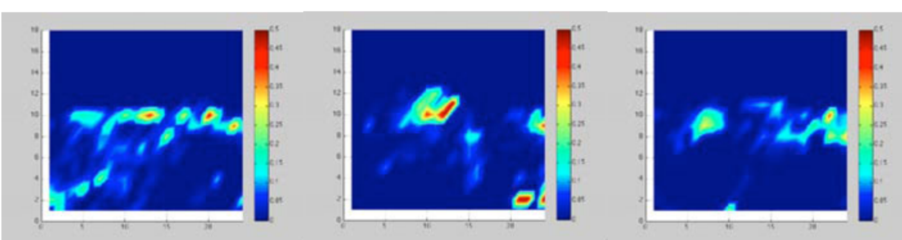
(a) 0-30 seconds (b) 30-60 seconds (c) 60-90 seconds

Fig. 12. Frequent distribution of staying time in case of regular shadow



(a) 0-30 seconds (b) 30-60 seconds (c) 60-90 seconds

Fig. 13. Frequent distribution of staying time in case of string-shaped shadow



(a) 0-30 seconds (b) 30-60 seconds (c) 60-90 seconds

Fig. 14. Frequent distribution of staying time in case of polygon shadow

in string-shaped shadow, at first, there are 5 to 6 red areas spread in parallel to the screen (Fig. 13(a)). Then it changed to the state where red areas are concentrated in the center (Fig. 13(b)) or to the state where red areas exist locally at either end of the screen as well as the center (Fig.13 (c)). Thus, we can see that there is temporal variation in position and number of the red area. That is, the children concentrated to one place while transformed shadow was projected, and then they spread in entire space from few clusters, repeating formation and collapse of groups. In addition, as Fig. 14(b) shows, in polygon shadow, groups were formed in perpendicular direction against the screen unlike the case observed in string-shaped shadow. The result indicates that group forming behavior differs depending on types of the shadow media.

4 Discussion

In this study, we newly developed a shadow media system, which could simultaneously lay out bodily expressions of around 20 people, to support activity of bodily expression which a group of children demonstrate. As a result, we found that, depending on difference in the shadow media, distinct images and bodily motions emerge within individual child. Regarding this, the authors assume that the shadow media works as image generator. That is, since the body and its shadow are inseparable, when we artificially transform the shadow, the body senses a feeling of gap. It is expected that this brings awareness to the body and consequently, images emerge. Reflecting expressions originated from the inseparable shadow media open up bodies of children and elicit potential images which their bodies possess. If that is the case, the result of this study implies that the way the shadow media elicits images varies depending on its type. In other words, depending on how we designs shadow media, we can elicit various kinds of bodily expressions.

This relates to how individual bodily expression improvisationally develops into that of groups through the shadow media in the bodily expression activity targeted at a group of children. That is because, it is necessary for a group of children demonstrating bodily expressions to unconsciously share what will likely to happen in the next moment of time. We think that, by sharing such image, each child plays each sharing role, and bodily expression, which is consistent with time and space, is realized. In that case, since the bodily shadow has a sense of equality, children can easily accept expressions of others, which create relationship of mutual dependence. Additionally, since the shadow media limits indefiniteness, co-creation of context and story, which gives a way to relate to each other, emerges. Also, we found that temporal and spatial phase change which forms a group could vary depending on what types of shadow media we use. For now, it is presumed that such a difference is produced because how to position the existence in the space differs depending on each shadow media type.

5 Summary

In this study, we aimed to utilize the shadow media system, which supports creation of bodily expression that we had developed, in the field of bodily expression activity in groups. Therefore, we developed a shadow media system which can be applicable to a group of 30 children at the maximum and used the system at Daini Sawarabi kindergarten (Okazaki, Japan). As a result, we have found that the shadow media system opens up individual body and improves bodily sense and that created expressions vary depending on the types of the shadow media. Also, we could observe how bodily expression of an individual expanded into an integrated bodily expression activity performed in groups. What's more, we discovered that this sort of bodily expression performed in groups by children formed a group and destroy it. Because of these, we think that the shadow media has effects which can not only draw out individual bodily expression through the relation with one's own shadow but also can support creation of bodily expression activity in groups by facilitating the connection between individuals.

Acknowledgment. This study was conducted under the project "Generation and Control Technology of Human-Entrained Embodied Media" and is supported by CREST of JST, Project Research "Principal of emergence for empathetic "Ba" and its applicability to communication technology" by RISE Waseda University, and the GCOE program "Global Robot Academia." We would like to express our sincere thanks to our graduate students, A. Kawai, T. Fukai for their assistants. Thanks should also go to for teachers and children of Daini Sawarabi kindergarten (Okazaki, Japan) for their cooperation for conducting experiments.

References

1. Miwa, Y., Itai, S., Watanabe, T., Iida, K., Nishi, H.: Shadow awareness: Bodily expression supporting system with use of artificial shadow. In: Jacko, J.A. (ed.) HCI International 2009. LNCS, vol. 5611, pp. 226–235. Springer, Heidelberg (2009)
2. Iida, K., Itai, S., Nishi, H., Miwa, Y.: Possibility for generation of body expression utilizing shadow media by measurement of body movement (in Japanese). Research Report of Human Interface Society 12(3), 71–74 (2008)
3. Okamoto, N.: Childhood -How children get the world?- (in Japanese), Iwanami-shinsho (2005)
4. Nishi, H., Noguchi, H., Yoshikawa, K., Hattori, M.: How can we apply three-dimensional computer graphics to creative physical expression? (in Japanese). In: Proceedings of Human Interface Symposium, pp. 137–140 (2008)
5. Izuhara, R., Nakamura, R., Miyajima, Y., Murasaki, K., Nagase, H.: Learning Support System of Physical Gymnastic Exercises for Blind (in Japanese), IEICE Technical Report, CQ2008-67, 2009-01, pp. 43–47 (2009)
6. Kawase, S., Osaki, A., Kaneko, T., Miwa, Y., Shiba, M., Tanaka, A.: Application of 3D Aerial Drawing System to Dance Therapy (in Japanese), IEICE Technical Report, HCS2007-67, 2008-03, pp. 21–24 (2008)

Virtual Interaction between Human and Fabric

Shigeru Inui, Akihiro Yoneyama, and Yosuke Horiba

Shinshu University, Faculty of Textile Science & Technology, Tokida 3-15-1,
Ueda, Nagano, 386-8567, Japan
{inui,10fa730c,horiba}@shinshu-u.ac.jp

Abstract. We tried to simulate human bodily movements to put on clothing. This is a combination of mechanical clothing simulation and measurement of bodily movements. A vest was modeled as virtual clothing because of its simple structure. Paper patterns were scanned and from the obtained images, virtual component panels of clothing were created. A model of clothing was constructed to seam virtually different component panels. The bodily movements to put on clothing were measured by a motion capture system. A virtual human model was moved according to the measured data of the movements. First, the shape of a vest in stable state hanging in the air was calculated, human body model was started to move, then the shape of the vest was calculated in every step of the movements. The movements to put on clothing were reproduced with the simulation.

Keywords: simulation, dressing activity, clothing, bodily movements, motion capture, BVH file format.

1 Introduction

The purpose of this study is to realize the simulation of dressing activity by bodily movements of virtual human model. The simulation consists of the mechanical simulation of clothing model and bodily movements of virtual human synthesis. We are aiming at the prediction and evaluation of the shape of clothing as the result of dressing activity and generation of the animation of dressing activity.

Simulations for cloth or clothing also have been studied [1-5]. Though there exists a fully geometric approach, mechanical approach is common for the purpose of clothing designing. Mass-spring model or continuum model has been utilized for cloth model of the simulation. Some of the clothing simulations are dynamic, and others are static. The practical application of the clothing simulation in the present situation is not designing or manufacturing but animation.

In conventional clothing simulation, prediction of the shape of clothing put on a human body has been focused, and the way to put on clothing has not been cared so much. There have been simulations of the change in shape of clothing accompanied by the bodily movements of virtual human, but the simulation of dressing activity is not be found. The way that virtual human puts on clothing is completely different [1-5] from the way we do everyday. In the simulations, it is usual that the panels of clothing are floating around a virtual human body in the initial state of the simulation,

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 418–424, 2011.

and then the panels are gradually approaching to the body and finally those are joined as clothing. Another approach is that the clothing model is created geometrically, and then it is changed into mechanical model [6].

The simulation to put on clothing has following advantages. The differences of bodily movements to put on clothing may cause the differences of the final state of clothing especially for under wear. It may become possible to predict the differences of the final state of clothing by realization of the simulation. Sometimes, persons with disabilities have difficulties to put on clothing. As the conditions of disabilities are different for each person, the clothing has to be designed in accordance with each person. It might be possible to design clothing for each person with disabilities if it becomes possible to predict the difficulties. Recently, care robots for elders have been studied [7]. In this field, the simulation might be applied to the control of robots that dress a care receiver.

2 Clothing Simulation

2.1 Formulation of Cloth Model

Clothing simulation here is based on the study of Okabe et al. [8,9]. Cloth is regarded as continuum and formulated based on theory of elasticity [10]. The strain of cloth model is determined from the deformation of cloth model. **r** is a position vector of a point on cloth surface in the three dimensional Cartesian coordinate system, and another coordinate system (u, v) the axes of which are fixed on cloth surface is considered. Tensile strain components are defined as functions of first order partial derivative of **r** with respect to u or v. Bending strain components are defined as functions of second order partial derivative of **r** with respect to u or v and **w** which is the unit vector normal to cloth surface.

The continuum cloth model was triangulated for numerical calculation, and strains were calculated form each triangular element. With mechanical properties of cloth, forces acting on each node of triangular elements are calculated from strains. The energy was minimized based on the force acting on each node of the triangular elements which is derivative of the energy. In the process of minimizing energy, nodes of triangular elements were moved until the energy minima was obtained.

2.2 Clothing Model

The processes to construct clothing model are shown in Figure 1. Paper patterns that are design drawings of component panels of clothing were scanned. The scanned images of the paper patterns were binarized, resized and noise was reduced using MATLAB. The virtual component panels of clothing were created from the data of the outlines of the scanned image. Each panel of the paper pattern was triangulated for numerical calculation. The panels were virtually sewn together and integrated as clothing model. Here, sewing means that corresponding nodes on seam in different panels were moved to the same point and identified as one node. The sewn clothing model is used for as the initial state of numerical calculation.

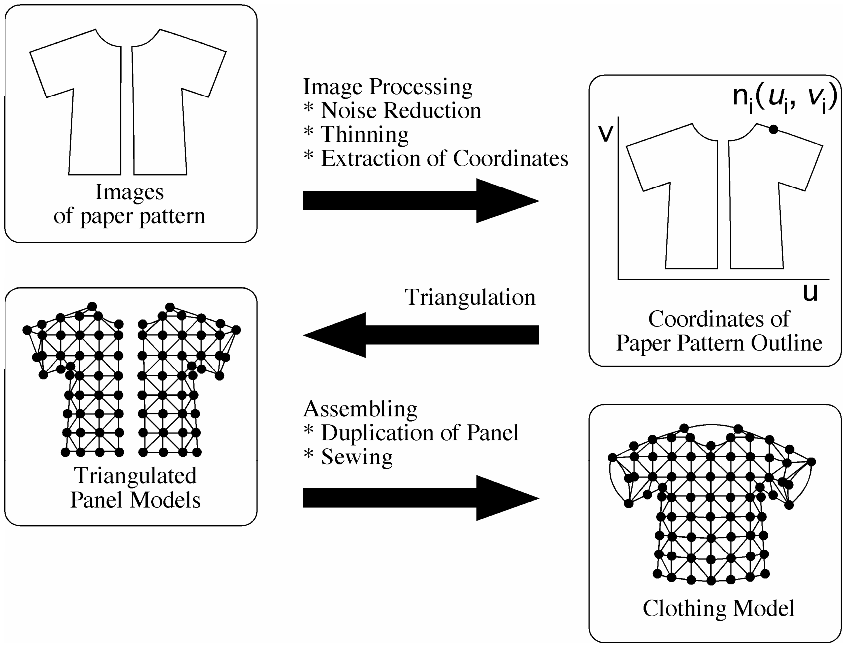


Fig. 1. Processes of constructing clothing model

2.3 Motion Capture

There would be several ways to generate bodily movements [11] for the simulation. The first way is to construct a mechanical human model, and to generate bodily movements based on geometrical and mechanical evaluation function [12]. The technique is now under development and it is difficult to generate complicated movements as dressing activity. The second way is to generate movements by interpolation of postures [13]. For this technique, various human postures are recorded in a database. To generate bodily movements, some intermediate postures in movements are defined, postures close to the intermediate postures from the database are extracted as key frames, and movements between the key frames are interpolated. Since bodily movements of dressing activity are complicated, it might be difficult to determine key frames. The third technique is to capture movements of dressing activity. This technique was adopted for this study because it may be the most feasible for our simulation.

An optical motion capture system "OptiTrack" of NaturalPoint Inc. was utilized to capture bodily movements of dressing activity. The system consists of a computer, eight infra red cameras and software "ARENA". When motion was captured, the eight cameras were arranged around a subject and the subject wore black tights on which reflective markers were attached. Ordinal clothing was not proper to use for measurement of dressing activity, since it sometimes block the markers while clothing was put on. The clothing for the capture was made of transparent fabric through which the markers could be seen. The bodily movements of a subject to put on clothing was captured and saved as a BVH formatted file.

2.4 Human Model

BVH format file assumes a skeletal structure. The structure consists of 16 joints, 5 end points (here we call those characteristic points) and 20 bones connect them. The skeletal structure is shown in Figure 2. In BVH file, the dimension of each skeleton was defined, and the rotational angles around each axis of each joint were recorded. The coordinate of each characteristic point was calculated from the rotational angles and the dimension of the skeleton of a human body model for each time step.

For simplicity of the program, the human body model parts such as trunk, arm, leg were defined as solids of simple shape such as sphere or cylinder. The structure of the human body model was defined based on the characteristic points.
Collision detection and reaction between the triangular elements of clothing model and the spheres or the cylinders of human body model were needed for the simulation. In this case, there are two types of collisions, one is a point and a cylinder and the other is a point and a sphere. When a collision between a point and a cylinder is detected, the point is moved to the surface of the cylinder. The same reaction is performed in the case of a collision between a point and a sphere.

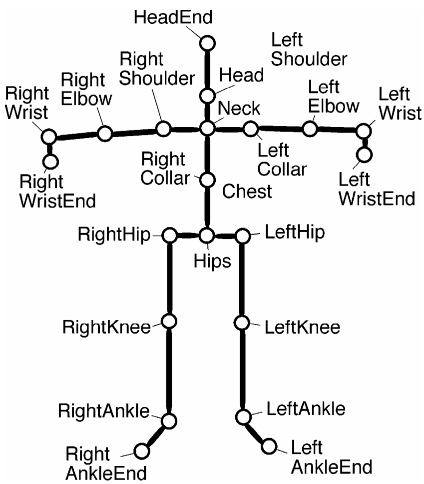


Fig. 2. The human model for the simulation

3 Simulation to Put on a Vest

3.1 Vest Model

Vest was selected as clothing for the simulation because it is easy to put on. Paper patterns of a vest were scanned, noise was reduced from the scanned images and then the images were binarized. The outline of the paper pattern was extracted as coordinate of points by thinning and edge detection of the scanned image. The virtual panels of a vest were generated form the outline data, and those were triangulated for numerical calculation. The nodes on seam line in different panel were moved to the same position to saw the seam line. Mechanical calculation was executed for the vest model to obtain an equilibrium state under the condition suspended at four points, and the shape was utilized for the initial state of the vest in the simulation of dressing activity.

3.2 Bodily Movements

We selected simple bodily movements for the simulation. The measured dressing activity was to put on a vest widely opened and suspended at two points on front collar (collar points) and two corner points of the front body (corner points) as shown in Figure 3. Each suspended point was fixed with string which could be disengaged by very weak force. The suspended position of the vest was determined from height of the subject and easiness to put on.

Following rules were set to definitely define dressing activity. The rules to put on a vest were,

(1) Begin with T shaped pause
(2) Hold a front corner point (either left or right)
(3) Put an arm opposite to the arm holding the corner point into an armhole
(4) Release the front corner point
(5) Hold a front corner point opposite to that of (2)
(6) Put an arm opposite to that of (3) into an armhole
(7) Release the other front corner point
(8) End with T shaped pause.

The number of steps of the measured bodily movements was 1,700 (100 steps per second), and the maximum movement of the characteristic point of the skeletal body model per one step was about 15 mm. Since the motion capture system utilized in this study could not record when and where the subject held clothing, some rules had to be

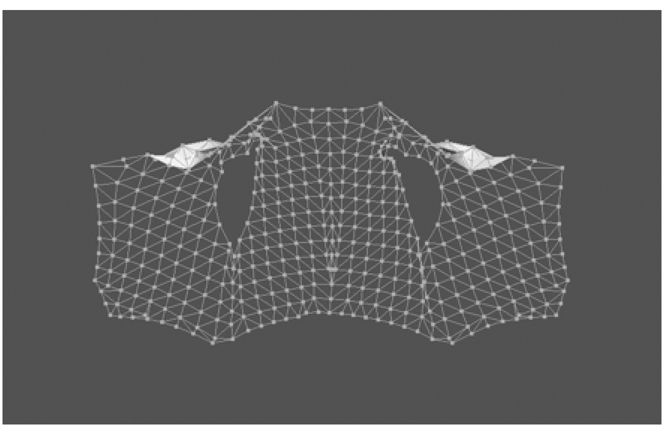


Fig. 3. The vest widely opened and suspended at two points on front collar and two corner points of the front body

set to execute the simulation. (1) When “LeftWrist” or “RightWrist”, the characteristic point of the skeletal human body model got close to a front corner point, “Wrist” was regarded holding the corner point. (2) When the direction of horizontal movement of the “Wrist” was changed, the front corner point and the collar point on the same side were regarded to be released.

4 Results and Discussion

The simulation to put on a vest consisted of two processes. The first step of the simulation, the initial suspended state of the vest was calculated. In the second step, the skeletal human body model was moved for one step of the measured movements, and the nodes of clothing were pushed out by the collision reaction. This is the initial state of this step and the minimum energy state of the clothing was calculated to obtain the shape of the clothing at the step. Then the human body model was moved to the posture of the next step, and the processes were iterated until the final step of the measured bodily mvement.

The dressing activity to put on a vest was able to simulate. The results of the simulation are shown in Figure 8a and 8b. In Figure 8a, Left side figures show both the human body model and the vest model and right side figures show only the vest model. The simulation might owe the easiness to put on a vest since there was no sleeve and small amount of displacement of the nodes of triangular element in one step.

Dressing activity is determined by the interaction between bodily movements and mechanical properties of clothing. Bodily movements may be affected by the properties of clothing such as light or heavy, stiff or soft. On the other hand, the shape of clothing may be different according to bodily movements or mechanical properties of the clothing. Rather soft fabric was used to make the real vest for the measurement of bodily movements of dressing activity. Mechanical properties of the cloth to make the vest for the motion capture was measured by the KES system which is usually used for measurement of fabric. Tensile property in warp and weft direction was 49.0 N/m, shear property was 49.0 N/m and bending property in warp and weft direction was 9.8 * 10^{-7} Nm^2/m. Mass of the cloth was 0.001 g/cm^2 and Poisson’s ratio of tensile and bending were set to zero. Those values were used for the mechanical calculation.

Since dressing activity is dynamic bodily movements, dynamic method may be utilized for the simulation. But there are some problems for dynamic method to predict the shape of clothing. It is difficult to determine damping property, and to verify the results of the dynamic movements. The result of numerical calculation varies according to integral algorithm. Because of the reasons above, a sequence of static mechanical calculation was utilized in the simulation. In the steady state, the results of static method are verified at least by visual judgment of the shape of cloth. To utilize static method for the simulation of dressing activity may correspond to put on clothing very slowly in the real world. In this case, we think it makes not so much difference of the shape of clothing when dressing activity is finished whether the calculation is dynamic or static.

Extensions of the simulation are expected. The first extension is the variety of clothing. This simulation method should be applied to clothing of more complex structure, for example with sleeves. The second extension is the way of dressing activity. Clothing should be picked up and put on by the human model itself. Improvements of the simulation and the measurement of bodily movements are needed for the extension.

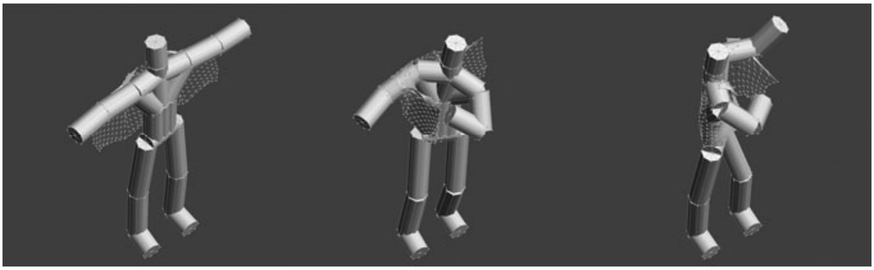


Fig. 4. The results of the simulation (Initial state, 600th step, 1200th step)

References

1. Sakaguchi, Y., Minoh, M., Ikeda, K.: PARTY: Physical Environment of Artificial Reality for Dress Simulation (1) – A Dynamically Deformable Model of Dress. Trans. Soc. of Electronics, Information and Communications: Graphics and CAD 54, 25–32 (1991)
2. Volino, P., M-Thalmann, N.: Virtual Clothing. Springer, Berlin (2000)
3. Breen, D.E., House, D.H., Getto, P.H.: A Physical-Based Particle Model of Woven Cloth. Visual Computer 8, 264–277 (1992)
4. Okabe, H., Imaoka, H., Tomiha, T., Niwaya, H.: Three Dimensional Apparel CAD System. Computer Graphics 26, 105–110 (1992)
5. Zong, Y., Xu, B.: Three-dimensional Garment Dressing Simulation. Textile Reseaarch Journal 79, 792–803 (2009)
6. Autodesk Inc. Maya cloth, `http://usa.autodesk.com`
7. Ikuta, K., Ishii, H., Nokata, M.: Safety Evaluation Method of Design and Control for Human-care Robots. International Journal of Robotics 22, 281–297 (2003)
8. Imaoka, H., Okabe, H., Akami, H., Shibuya, A., Aisaka, N.: Analysis of Deformations in Textile Fabric. Sen'i-Gakkaishi 44, 217–228 (1988)
9. Imaoka, H., Okabe, H., Akami, H., Shibuya, A., Aisaka, N., Tomiha, T., Yamada, M.: Prediction of Three-Dimensional Shapes of Garments from Two-Dimensional Paper Patterns. Sen'i-Gakkaishi 45, 420–434 (1989)
10. Green, A.E., Zerna, W.: Theoretical Elasticity. Dover Publications, Inc., New York (1992)
11. Xia, S., Wang, Z.: Recent Advances on Virtual Human Synthesis. Science in China Series F: Information Science 52, 741–757 (2009)
12. Liu, Z., Gortler, S.J., Cohen, M.F.: Hierarchical Spacetime Control. In: Proceedings of SIGRAPH 1994, pp. 35–42 (1994)
13. McCann, J., Pollard, N.S., Srinivasa, S.: Physics-Based Motion Retiming. In: Proceedings of the Eurographics/ACM SIGRAPH Symposium on Computer Animation 2006, pp. 205–214 (2006)

Hand Gesture-Based Manipulation of a Personalized Avatar Robot in Remote Communication

Teruaki Ito

The Institute of Technolgy and Science,
The University of Tokushima,
2-1 Minami-Josanjima, Tokushima 770-8506, Japan
ito@me.tokushima-u.ac.jp

Abstract. This study proposes an idea of hand gesture-based manipulation of an personalized avatar robot for video conference and presents the on-going work for this study. To achieve this goal, a remote controlled communication robot called Collabo-Bot is under development, which has face projection, maneuvering and manipulation functions to improve the better communication over the network. One of the challenging ideas in this study is to personalize the communication robot using face projection to represent the operator. By doing this, the remote operating person who controls the robot enables the video conference attendees at the local site to feel the existence of the operator much stronger than a regular video conference without using a robot, or even than a typical robot-based conference. In addition to that, hand gesture-based manipulation is under study to control the avatar robot. This paper overviews the design and manufacturing of Collabo-Bot and presents some experimental results on the implementation of this hand gesture-based manipulation.

Keywords: Network-based robot, hand gesture-based manipulation, robotic arm manipulation, human interface, face projection.

1 Introduction

Network-based communication is getting popular according as the development of video conference systems (Abowdm 2000). Even though the remote communication is possible over the network using these systems, it is generally difficult to feel the existence of those participants who are physically located on a remote site as opposed to a face-to-face meeting (Greenberg 1996). Under these circumstances, humanoid robots attract attention as an avatar which works remotely in favor of the operator. Using a humanoid robot for remote communication with multimodal information, the experimental results in several studies show the effectiveness of these remote controlled robots in communication. For example, GestureMan (Kuzuoka 2000), wakamaru (MHI), Giraffe (HeadThere) belong to this category. These robots provide basic function to suppose distance communication using several critical functions such as face image display of the operator, drivability to move around, tele-manipulation on remote objects as well as basic communication functions including talk/listen/see. Even a

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 425–434, 2011.

smaller robot is now available on the market for these purposes as well. For example, Rovio (WowWee) is a Wi-Fi enabled mobile webcam robot that lets the user view and interact with its environment through streaming video and audio. There robots could provide an existence of the operator in the remote site and even enables some kinds of tele-operating tasks from distance. However, there are still a gap between robot-based video conferences/meetings and face-to-face ones, which is an open issue to be solved for better communication.

First, this paper describes the background of the communication robot in the applications such as remote instruction (Singh 2007), distance learning class work (Ogata 2007), etc. Then this paper presents the remote controlled communication robot in this study, called Collabo-Bot, in terms of its design and implementation. Presenting the experimental implementation on the hand gesture-based manipulation, this paper discusses the feasibility of this manipulation in video conferencing purposes.

2 Research Background

Internet-based video conference is getting popular these days thanks to the advanced information technology. However, as opposed to the face-to-face meeting, video conference does not always work to support good communication because of the feeling of non-existence of the remote participants (Sugimoto 2004). One of the solutions to this issue would be to use a remote controlled robot to represent the operator, which is stated at the introduction section. As for the operator side, the robot could take care of the task which is supposed to achieve on a remote site by way of tele-manipulation of the operator. As for the participant side, the remote controlled robot provides an existence of the operator as well as performing a tele-operation task thanks to the fact that a physical object explicitly exists there. For both sides, enhancement of mutual communication could be expected not only by using sound/audio media, but also by carrying out interaction with the remote controlled robot. However, it is not clear how to use the robot in an effective way in video conference. In order to study the feasibility of the effectiveness of the robot as expected, experimental evaluation is required. The challenge of this study is not simply to use a communication robot, but also to personalize the communication robot so that the robot more appropriately represents the operator on the remote site. The idea behind this is to implement a face projection function on the robot. However, it is not clear how effective the face projection is. Therefore, experimental study is also required to show the feasibility of the proposed idea.

This research is developing a remote controlled communication robot called Collabo-Bot, implementing some basic functions for remote controlled robot. In this research, the robot could not merely work as a slaved avatar of the operator but also work as a partner which represents the operator on a remote site. To do so, this study focuses on an effect of face projection of the operator on the robot, to clarify how to provide an atmosphere of the operator on a remote site. In this way, this study will pursue to find a clue to clarify the effective use of face projection for further enhancement of robot-based remote communication as the research goal.

3 Collabo-Bot: Its Design

3.1 Overview of Collabo-Bot Design

To achieve the goal of remote communication support, a remote controlled communication robot called Collabo-Bot is under development. Collabo-Bot is composed of several basic units; such as maneuvering module, projection module, manipulation module, communication module, and conference module. All of these modules are controlled by a Laptop PC with WiFi connection module installed inside of Collabo-Bot. This section covers these basic units and presents the basic design of Collabo-Bot.

Figure 1 shows the overview of Collabo-Bot components, including five modules and three units to configure the robot. The five modules include internal processing modules, whereas the three units include the basic physical framework to configure the robot. Considering the use of Collabo-Bot in a meeting or in a class, the size of the robot is designed as 1500mm height and 600mm width as shown in Figure 1.

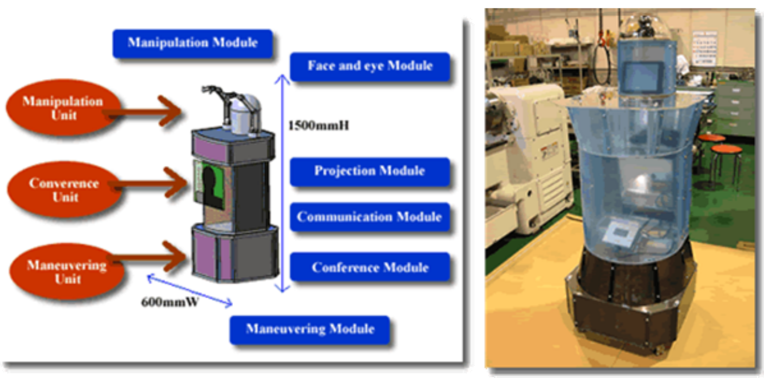


Fig. 1. Basic components of Collabo-Bot and its overview photo

3.2 Maneuvering Module with Four Omni-Wheels

The driving module of the robot is composed of four sets of omni-wheel (Mfg by Tosadenshi) with geared DC motor (Tamiya), which are controlled by a servo motor controller (PhidgetMotorControl LV). Tamiya's DC motor is operated by 7.2 V batteries and the maximum load of the driving module is 30Kg, which meets the requirements of the design of Collabo-Bot. The details will be given in section 4.

3.3 Projection Module for Face Projection

The projection module is installed in the middle portion of Collabo-Bot. It is designed not only to show the face of a remote operator, but also to show presentation slides/images controlled by the remote operator. Since the size of the middle part of the robot body is designed as 500mm width, 600mm height and 500mm depth, straight projection to the screen exceeds the available projection distance. Therefore, a mirror projection was designed. The details of the design and its manufacturing will be shown in section 4.

3.4 Face and Eye Module with Web Camera and a Small LCD

An eight inch LCD monitor (Century) is installed to show the face of the remote operator inside of the head portion of the robot attached at the top plate panel. A web camera (Logicool) is installed above the monitor to work as an eye of the robot, as shown in Figure 2. Even though the face projection is the critical part of Collabo-Bot, this LCD monitor is installed to compare the feasibility of remote communication support.

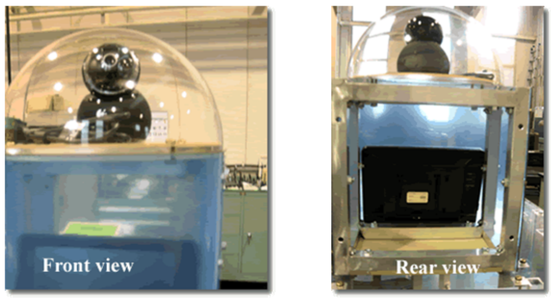


Fig. 2. Face and eye modules

3.5 Manipulation Module with Robotic Arms

In front of the head unit, two sets of robotic arm (Lynx 5 robotic arm) with motor controller are installed on the top panel of Collabo-Bot. These arms are controlled by a remote operator during video conference; for example, raising a hand to pose a question during conversation, pre-defined several hand gestures to support communication, manipulation with the arms on a remote object, etc. Figure 3 shows the photo of the robotic arms and its CAD modeling images. As opposed to a standard video conference based on audio/image communication, it would be possible to perform these remote manipulations using Collabo-Bot.

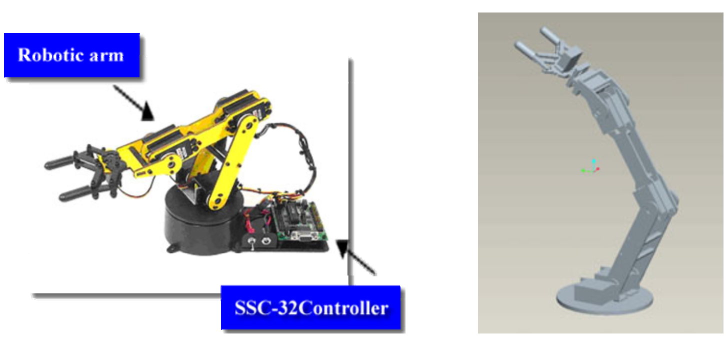


Fig. 3. Manipulation module: robotic arms image and its CAD model

3.6 Communication Module and Conference Module

There are various video conference software available today. Among these, this research adopted Skype, which allows users to make voice and video calls over the internet. Skype is installed on the Labtop PC inside Collabo-Bot to make video communication.

Skype chat allows both calling a single user and a conference calling. Since Collabo-Bot is pursuing conference calling, a speaker phone (Sanwa Supply) is installed inside of the robot body so that multiple person could chat with the remote users through the robot.

4 Collabo-Bot: Its Manufacturing

4.1 Overview of Collabo-Bot Manufacturing

Collabo-Bot is under development at the University of Tokushima based on the design ideas presented in the previous section. This section presents how the manufacturing is conducted to build Collabo-Bot.

4.2 Components of Materials and Assembly Units

Figure 1 in the previous section showed that Collabo-Bot is composed of three core units, which are the maneuvering unit at the bottom portion of the robot, the conference unit in the middle portion of the robot, and the manipulation unit at the top portion of the robot. These units are under manufacturing respectively, and will be assembled to configure the robot in the final stage of manufacturing. In this section, the current status of manufacturing will be covered briefly.

The internal structure of maneuvering unit is made with a frame structure of square aluminum pipes and acrylic inner plates, which is covered by a set of polycarbonate panels. Inside of this unit, several key components are installed; such as maneuvering module, controller board, battery unit, and connecting terminals. The laptop PC for control is also installed inside of this unit. Figure 4 shows the CAD image of this unit as well as the manufactured unit. Control program for remote operation is also under development to be used during video conference.

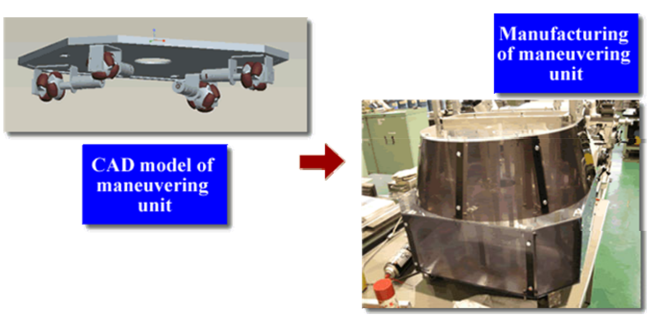
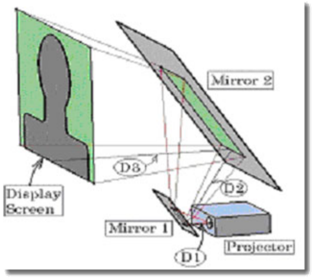


Fig. 4. Manipulation module design and its manufacturing



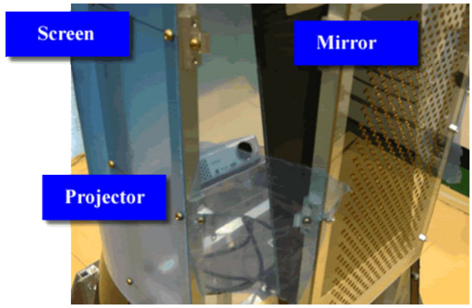


Fig. 5. Projection unit design and its manufacturing

The conference unit, which is placed in the middle of the robot, is also built on a square aluminum pipe with acrylic plate inner structure. The front panel is installed with a transparent polycarbonate sheet covered with a rear projection screen. To keep the appropriate projection distance between the projector and the screen, a combination of two mirrors is used as explained in the previous section. Figure 5 shows the design of rear projection unit and its manufacturing. The original screen in the prototype (Ito 2004 & 2005) is based on a white colored polycarbonate panel. For better view of the projection image, a combination of transparent polycarbonate plate and a special rear projection sheet is used. However, the shape of the projection panel of the robot is not flat but has a curved surface. Therefore, further experiments are scheduled to achieve the high quality of projection.

The third unit of Collabo-Bot, or a manipulation unit, at the top is built on an aluminum frame with several acrylic panels attached with a dorm cap over it. This unit also includes a small LCD display, a remote controlled web camera, and the two sets of robotic arms with a gripper. All of these are assembled and configured to build this unit. Especially, the user interface for the arms is very critical in order to provide a user-friendly and intuitive interaction with the robot.

5 Hand Gesture-Based Manipulation of Collabo-Bot

Hand gesture plays a significant role in face-to-face communication. It has the potential to express one's feeling or to describe one's situation in an intuitive way as a non-verbal language. This study assumes that a hand gesture could be used as an interface for a robot arm manipulation during the video conference and reviewed how to use the hand gesture for this purpose. This section presents a method for gesture-based arm control in this study and compares it with other forms of controlling operation; including joystick-based manipulation and voice-command-based manipulation (Ogawa et al. 2007).

Implementation of hand gesture for arm manipulation has been achieved in several steps, which is also presented in this section. Figure 6 shows the basic components of robotic arm manipulation for the gesture manipulation experiment. The components include P5 data glove, Lynx 6 robotic arm, SSC-32 Controller, StrokeIt and RIOS software.

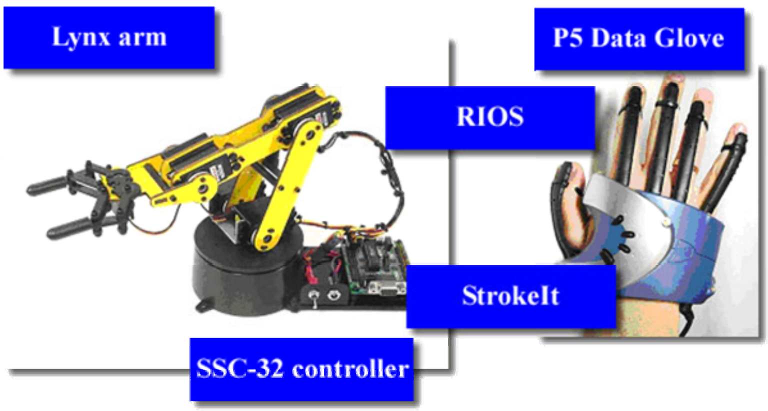


Fig. 6. Components for hand gesture-based robot arm manipulation

5.1 Experimental Description

By using the RIOS software, several sets of arm manipulation operation have been made and their data were saved under different projects in play module. In other words, each complete manipulating operation (for example, grasping an object in front of the arm and put it in a definite location) is composed of several sequences, and they were saved as manipulation projects.

5.2 How to Create a Gesture Using P5 Data Glove

While the P5 glove is in active mode, it is difficult to recognize the starting point of a gesture. The major reason of this is that the P5 glove continuously sends its position signal to the receptor tower. To recognize the starting point of a gesture, this study uses a combination of an index finger and a thumb finger. When an index finger and a thumb finger touch each other, a gesture recognition process starts until it finishes with the release of these two fingers. In other words, the gesture recognition process stops when these fingers are detached.

5.3 Gesture-Based Arm Manipulation

Using the P5 data glove, several simple gestures were registered. A gesture recognition software, or called StrokeIt, is used to identify these gestures. A new gesture is defined in learning mode of StrokIt. An application program of windows or group of programs can be configured for special use with StrokeIt. When P5 glove function properly makes expected gesture, StrokeIt recognizes it and initiate a stored movement project in RIOS.

5.4 Gesture Recognition and Successful Manipulation Rate

Not every gesture by the P5 glove was successfully recognized. One of the reason why is that a human hand often quivers to some degrees when it is placed in a free space. The starting point of a gesture is very critical, because the P5 glove always needs to be placed within three to four feet range of the receptor tower. Sometime the receptor tower lost the signal from the P5 glove, and an expected dynamic gesture was not performed.

5.5 Discussions on Pros and Cons of Gesture-Based Interactions

a. A distinguishing feature of gesture based command would be that it could be terse and powerful for interaction. However, the experiments in this study showed that the major obstacle for gesture interaction was associated with the use of P5 data glove. The P5 data glove needs to attach to sensor components and wires, which limit the user's freedom of movement. It also limits the range in space where it can be operated.

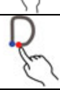
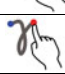
b. Making-gesture causes hand-fatigue very easily. During mice or joystick interaction, user's hand is supported by a desk or a table which minimizes the fatigue.

c. In the gesture based control, the operator must know a well-defined set of gestures to initiate a particular movement of the robotic arm. In keyboards or joysticks, typical key functions are well known and used; for example, up, down, left or right arrow keys in keyboards or joysticks are well understood as the functions to the up, down, left or right movements, respectively. However, there is no universal set of gestures in order to operate these functions.

d. A Joystick is used to control the robotic arm and to do the same kind of manipulating operation done by hand gesture. In joystick, it is quite difficult to make a precise operation. In order to grasp an object and to place the robot hand gripper to the right position, operator should closely observe the manipulation. Since joystick is a common controlling device, its installing process and smoothness of operation is very good.

e. Voice-command-based arm control is another form of interactive arm manipulation, especially when an operator knows the necessary commands to manipulate the robotic arm. For executing the operation, voice command software (e-speaking) is used. RIOS is configured to operate through voice command. In the experiment, while the hand gesture based manipulation method suffered form the incapability of producing gesture (by P5 glove), the voice command based control had no such obstacles.

Table 1. Several examples of gesture-making attempts and successful arm manipulation rate

Gesture type	**No. of gesture making attempt**	**No. of successful recognition by StrokeIt**	**No. of successful manipulation of robot arm**	**Successful manipulation rate**
	10	6	6	60%
	10	5	5	50%
	10	6	6	60%
	10	4	4	40%

6 Concluding Remarks

This paper focuses on the use of communication robot to increase the feeling of existence of a remote participant in video conference over the network. The idea to use one of these communication robots is not new. However, personalization of the robot using the idea of face projection and hand gesture-based manipulation are the major challenges pursued in this study.

The design of Collabo-Bot is almost finished and its manufacturing is under way. The idea of rear projection inside of Collabo-Bot has been reviewed as feasible from the experimental prototype evaluation. Manipulation program of robot arms with grippers, and maneuvering operation of the robot are also under development. Based on these achievements, the next step of manufacturing is to assemble all of these units and configure them to build Collabo-Bot. More detailed description of manufacturing status of the robot will be presented in the conference. The experiments on the effect of face projection and hand gesture-based manipulation will follow after that.

Acknowledgement. This research is partly supported by the Grant of Industry Academic Joint Researcher Fund of the University of Tokushima in 2009. The author would like to thank Mr. Tariq Arif Mohamad at the Collaborative Engineering Lab at the University of Tokushima for his help in modeling of the robot. The author would also like to thank Mr. Junji Tamatani and his staffs at the Center for Technical Support of the University of Tokushima for their cooperation in manufacturing. The author would like to thank Prof.Hiroaki Ogata and his staffs of the University of Tokushima for their comments on computer-based learning using a communication robot.

References

1. Abowdm, D.G., Mynatt, D.E.: Charting past, present, and future research in ubiquitous computing. ACM Transactions on Computer-Human Interaction (TOCHI) 7(1), 29–58 (2000)
2. e-Speaking, `http://www.e-speaking.com`
3. Giraffe, `http://www.headthere.com`
4. Greenberg, S.: Peepholes: low cost awareness of one's community. In: Conference Companion on Human Factors in Computing Systems: Common Ground, Vancouver, British Columbia, Canada, pp. 206–207 (1996)
5. Ito, T.: An approach of whiteboard-based interface for collaborative environment. In: 11th European Concurrent Engineering Conf., Hasselt, Belgium, pp. 71–75 (2004)
6. Ito, T.: An Approach Towards Collaborative Communication. In: The 12th ISPE International Conference on Concurrent Engineering, Fort Worth, Texas, USA, pp. 247–251 (2005)
7. Kuzuoka, H., Oyama, S., Yamazaki, K., Suzuki, K., Mitsuishi, M.: GestureMan: a mobile robot that embodies a remote instructor's actions. In: Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work, Philadelphia, PA, USA, pp. 155–162 (2000)
8. Lynxmotion, `http://www.lynxmotion.com`

9. Ogata, H., Terai, K., Shinto, T., Bando, W., Yano, Y.: Supporting Awareness in Real-Time Distance Learning with a Real-World Avatar. In: Proc. of the 2007 Conf. on Supporting Learning Flow Through Integrative Technologies, vol. 162, pp. 425–432. IOS Press, Amsterdam (2007)
10. Ogawa, H., Watanabe, T.: InterRobot: a speech driven embodied interaction robot. In: Proceedings of the 9th IEEE International Workshop on Robot and Human Interactive Communication, RO-MAN 2000, September 27-29, pp. 322–327 (2000)
11. P5 glove, `http://www.vrealities.com`
12. Phidgets, `http://www.phidget.com`
13. RIOS, `http://www.lynxmotion.com`
14. Rovio, `http://www.wowwee.com`
15. Sugimoto, M., Hosoi, K., Hashizume, H.: Caretta: a system for supporting face-to-face collaboration by integrating personal and shared spaces. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Vienna, Austria, pp. 41–48 (2004)
16. Singh, D., Abu Bakar, Z.: Wireless implementation of a mobile learning environment based on students' expectations. International Journal of Mobile Learning and Organisation 1(2), 198–215 (2007)
17. Skype, `http://www.skype.com`
18. StrokeIt, `http://www.tcbmi.com/strokeit/`
19. Wakamaru, `http://www.mhi.co.jp/kobe/wakamaru/english/`

Vector Keyboard for Android Platform-Based Devices

Martin Klima and Pavel Slavik

Czech Technical University in Prague, Faculty of Electrical Engineering,
Karlovo nam. 13, Prague, Czech Republic
{xklima,slavik}@fel.cvut.cz

Abstract. This paper introduces a Vector keyboard for mobile devices with a touchscreen. Typing is produced by simple strokes sourcing from one of four blocks of letters. A user study was made comparing this keyboard with standard QWERTY, ABC and a Mobile keyboards. The study resulted in a comparable performance in terms of speed and typing with all the named layouts but QWERTY. The user preference and subjective evaluation is in contrast to the performance measured preferring the Vector keyboard over ABC and Mobile.

Keywords: vector keyboard, virtual keyboard, touch screen, PDA, QWERTY, ABC.

1 Motivation

The goal of this research is to introduce a new type of keyboard design for touch screen mobile devices, which would be immune to disturbing influence of surrounding when typing in a mobile environment.

There are several commonly used virtual keyboards designed for either mobile phones or handheld devices in general whose performance is degraded when using on small screens. An example of such a keyboard is the QWERTY keyboard, which is present on all smart-phone operating system like Android, iPhone OS, Windows Mobile and others. Such keyboard, when displayed on a small screen, does not scale well with degreasing screen size and its performance drops dramatically. Therefore we designed a keyboard layout which scales well and which should perform better on touch-screen mobile devices. This study is about verification and evaluation of usability of our modified keyboard and about objective performance evaluation and subjective user experience results. We focused solely on comparison of virtual keyboards that make it possible to enter text using fingers only (e.g. no stylus is necessary). No prediction algorithms was used with any of the keyboards measured since we wanted to evaluate the core qualities of the keyboard, not the quality of prediction algorithm used.

2 State of Art

Our research is primarily based on our previous work [1] in this area, which introduced the first version of a vector keyboard for mobile devices. This design used a

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 435–442, 2011.

combination of tapping on the touch screen and drawing a vector. Each action (tap or draw) resulted in typing one character.

Fig. 1 shows the layout of the keyboard which consisted of three blocks of alpha-numerical characters, each blocks organized in a matrix of 3x3 characters. The blocks can be exchanged for a different set of three blocks by either pressing the shift button (upper left) or by pressing the symbol button (lower left).

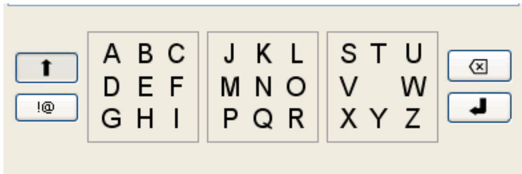


Fig. 1. The layout of vector keyboard designed in [1]

Typing on the keyboard was done by two principles.

- The first and major one was drawing a vector with a starting point within one of the three blocks and heading in a direction which is parallel to a vector starting from the center of the block heading to the character to be typed. The principle is demonstrated in Fig. 2 on typing letter "a". Please notice, that the starting point of the vector can be anywhere in the area. Such design makes it easy to locate the relatively large block.
- The second one was tapping anywhere in the area which resulted in typing the character in the middle of the block. An example is typing letter "n" in Fig. 2.

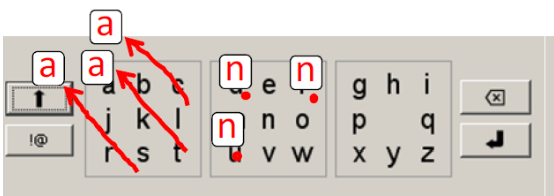


Fig. 2. Typing letters **a** and **n**

The major problem of the approach was that the user often confused tapping and drawing very which increased the error rate. The study as introduced in [1] was carried out on an ultra mobile PC with a resistive display which did not correspond to regular mobile phones in terms of display size, touch screen sensitivity nor the size and weight of the device.

Besides the above introduced technique, there are methods that employ gestures for typing individual characters or whole words. A nice overview is given in Jochen Rick [3] where a number of layouts for virtual keyboards is summarized. Most of them are a variation to tapping QWERTY keyboards and thus not differing from the well known layout. An interesting approach is the Cirrin [4] where the letters are organized in a one circle and word is written by drawing a stroke through all the letters. This approach was later improved by Cechanowicz et al [5]. This approach uses one stroke to type multiple letters at once. Since our primary focus is on a mobile environment where the user might be disturbed in producing long and complicated gestures, we do not find this method suitable. An interesting approach is Quikwriting [9][10] where individual letters are written by drawing gestures starting and ending in the same point. The path of the stroke must follow complicated rules in order to distinguish which of 8 letters in one region should be typed.

The most well know virtual keyboard lately is the Swype keyboard which has the ability to degrade to a standard QWERTY keyboard but can also offer a fast gesture based word typing. Since this method is based on dictionary and language model based algorithm, it is out of scope of our study (same as other well known methods including Dasher).

3 Problem Description

There are several major problems of the existing vector keyboard designed in [1]. Firstly the layout of the three blocks is not suitable for mobile devices like smart phones with their display size (portrait orientation). The three blocks simply do not fit to the display apart from the controlling elements that make the space indigence yet more serious.

Secondly the nature of the typing and the controlling of the other elements - delete, shift, enter, symbols is different. While the typing is done by strokes, the other named actions are made by clicking (tapping). This may confuse the user.

Thirdly there is an internal inconsistency in the way the individual characters are typed. While some are typed using strokes, others are typed using tapping. This fact was identified as a problem in the previous study and users did complain about this.
Our aim was to:

- Redesign the vector keyboard to be suitable for the commonly used smart phones using Android platform.
- Make a qualitative usability study comparing the most common methods of text inputs. These methods should not be based on a language model or dictionary.

4 Solution

Our keyboard layout consists of four blocks, each containing eight different letters (see Fig. 3). A letter is typed by drawing a vector (stroke) from one of the blocks in a direction which is parallel to vector directing from the center of the block to the letter on its border. In contrast to the previous work, there is no letter in the middle of the rectangle, no tapping is ever needed.

All special actions needed for managing the texts are distributed in the blocks. This covers the delete, enter, shift, space and other alphanumeric characters.

In the center of the four blocks, there is a special sliding button for changing the set of typed letters. By dragging this button to the left or right, it is possible to switch to a different set of letters. Dragging the button to the bottom hides the keyboard from the visible screen area.

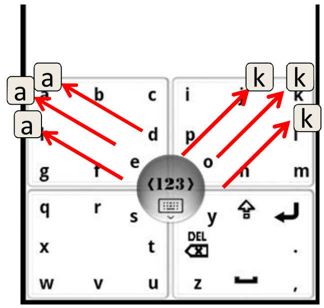


Fig. 3. Vector keyboard layout and usage

The study. We designed our study based on a hypothesis that the keyboard is more error prone than QWERTY and ABC layout keyboard of a comparable size and that the total typing speed will be faster and thus the vector keyboard more efficient. Prior to the study we had two users using the keyboard for a two weeks period who reported better user experience compared to the QWERTY. The reported positive experience was prior to all based on situations where it was not possible to use any prediction algorithm.

The study consisted of comparison of four different keyboard types, namely QWERTY, ABC, 12 buttons cell phone like virtual keyboard (a representative of numerical keyboard of a traditional cell phone) and our vector keyboard. We collected data from 18 users who typed equally difficult texts on each keyboard type. All of the keyboard had built-in logging mechanism which recorded all user actions and timings.

We used four different texts originating from a web site informing about latest financial news. All four of them had equal length of 100 words, equal difficulty and equal frequency of special characters like numbers, parenthesis or quotation marks.

The texts were dictated so that the user did not have to spend time thinking about any grammar and spelling issues. Keyboards and texts were used in a random order, users made pauses between individual typing for at least 10 minutes.

Device used for the test was a HTC Hero with Android OS. Users were sitting during the tests. The style of typing was up to the user whether he/she used one or two hands, left or right.

The users were of age from 16 to 28 (mean 22 and std. dev. 2.76), 6 users had previous experience with typing texts on a touch screen smart phone using QWERTY virtual keyboard, 12 users did not have any previous experience with smart phones. All users were experienced in using regular mobile phones and PC with QWERTY or QWERTZ keyboard.

The test covered the first experience with the vector keyboard, no training except of introduction to the principles of typing was given to the users.

5 Results

We used several metrics for evaluation of the typing performance.

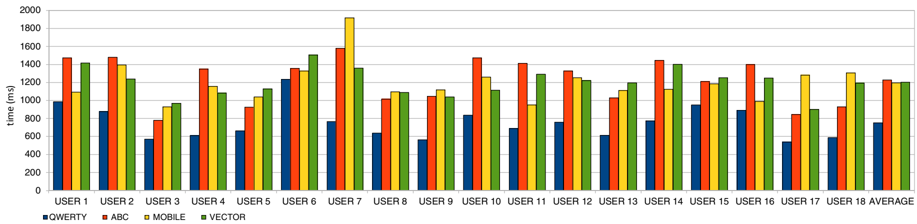


Fig. 4. Raw typing speed including typos and spaces (lower is better)

First a typing speed of a single letter was evaluated with no filtration on typos, pauses between words and first letters in words and is presented in Fig. 4.

QWERTY in this test performed best, probably thanks to the fact that users were familiar with its layout from the PC. The other three keyboards performed equally well in average.

Table 1 shows relative speed compared to QWERTY keyboard as a reference.

Table 1. Relative keyboards speed with no regard to typos (average)

QWERTY	ABC	MOBILE	VECTOR
100%	163%	159%	160%

Second evaluation was performed on a data set, where we extracted only such letters that were inside a word, e.g. starting from the second letter in the word. Spaces and single letter words as well as typos were removed. The reason for such extraction was our observation that users needed extra time for preparation for typing between words rather than in the middle of a word. By excluding the first letter and the spaces from this statistics we subtracted the mental preparation phase for the whole word from the average time needed to type a single letter.

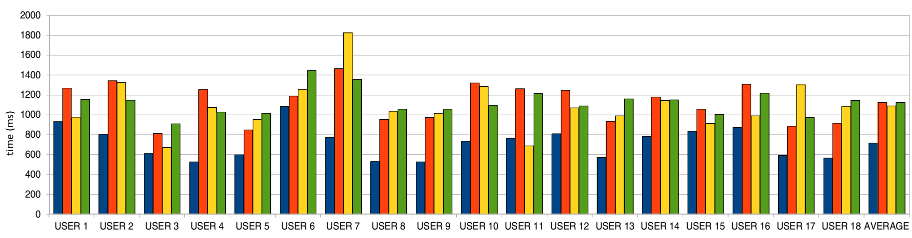


Fig. 5. Typing speed with no typos, spaces and first letters

The relative slowdown of the virtual keyboards compared to the QWERTY is shown in Table 2.

Table 3 shows the difference in results for the named methods. All keyboards exhibit faster times when the major mental preparation phase for typing is excluded. It is a very interesting observation - the QWERTY is the most well know layout and the mental preparation plays significantly smaller role. It is difficult to extract the mental preparation phase between individual characters types in the first experience study. It could be observed in a longer study as a typing speed improvement (learning).

Table 2. Relative keyboard speed with regard to typos (average)

QWERTY	ABC	MOBILE	VECTOR
100%	156%	152%	156%

Table 3. Speed of typing method evaluation

QWERTY	ABC	MOBILE	VECTOR
5%	9%	10%	7%

In contrast to our expectation, the vector keyboard did not perform significantly different from the ABC and Mobile layout. The evaluation of the errors, e.g. typos was done by measuring all initial typos made in a word. In case the user made a typo and continued typing without notifying, than noticed the typo, deleted all characters following the typo and typed everything again, we considered this just as one typo. This way we excluded false typos caused by the nature of individual users.

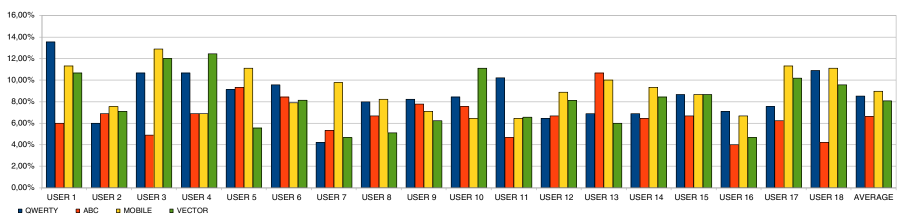


Fig. 6. Error rates

The error rate produced during typing on individual keyboards is displayed in Fig. 6. In average the vector keyboard did not perform significantly better than other keyboards, which did not prove our hypothesis. It is that some users performed very well compared to QWERTY (1,6,8,9,11,16) but for some users the typing philosophy of the vector keyboard was extremely unusual (3,4,10,17) and did not reach a good performance during the short testing period.

Finally we did a subjective evaluation of the keyboards where the users assigned to 1 to 4 (best) points to each keyboard expressing its speed, error proneness and usability. The sums of the points is shown in Fig. 7.

Surprisingly, the results are not in a correlation with the performance measured. Users definitely prefer the QWERTY layout, which also performed best. Despite the fact that the ABC keyboard was performing equally well as the Mobile and Vector

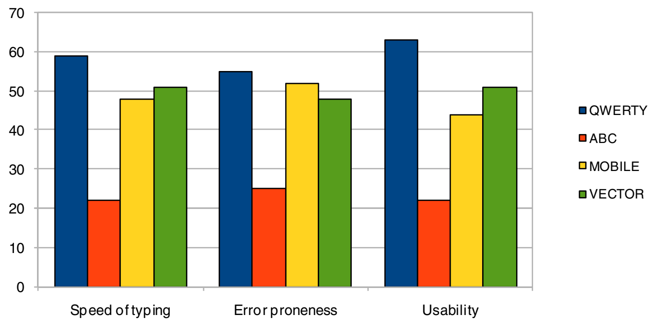


Fig. 7. Subjective keyboard evaluation

keyboard, it was evaluated significantly worse. Subjectively the usability of the Vector keyboard received the second highest score.

As a subjective observation we identified a problem with typing letters in the lower right block of letters when holding the device by right hand and producing stroke in the direction to the right and right bottom. Similarly the same issue arises when holding the device by left hand and typing letters located in the left or left bottom area. This problem is caused by the anatomy of human hand – thumb. We will consider this issue in the next versions and will adapt the system of coordinates to reflect the natural move capabilities of human hand.

6 Conclusion

We measured the real and subjective performance of the Vector keyboard implemented on a smart phone with Android platform. We compared the Vector keyboard with other commonly used layouts (QWERTY, ABC, Mobile) with no typing prediction. Our initial hypothesis that is will naturally perform better than a QWERTY keyboard did not prove true. The keyboard is performing similarly well as the ABC and Mobile keyboards while the QWERTY wins. One of the interesting observation was the fact, that the cognitive load (mental preparation) is significantly higher by all the ABC, Mobile and Vector keyboards, almost a double the value of the QWERTY. The performance is in a correlation with this finding.

The subjective evaluation of speed, error proneness and usability shows that users give priority to the Vector keyboard over ABC and Mobile.

As a possible improvement and a future work we now consider implementation of a language model and a prediction. The user types one letter per stroke and the stroke hits ideally one letter. In the worst case we can consider also the neighboring letters. The prediction can take advantage of this limited set of letter that come in question.

Acknowledgment. This research has been partially supported by the MSMT under the research program MSM 6840770014. This research has been partially supported by the MSMT under the research program LC-06008 (Center for Computer Graphics).

References

1. Klima, M., Slovacek, V.: Vector Keyboard for Touch Screen Devices. In: Proc HCII 2009: Ergonomics and Health Aspects of Work with Computers, pp. 250–256. Springer, Heidelberg (2009)
2. Po, L.M., Wong, C.K., Ting, C.W., Ng, K.H., Wong, K.M., Cheung, C.H., Cheung, K.W.: Dynamic Candidate Keypad for Stroke-based Chinese Input Method on Touchscreen Devices. Accpeted for Publication in 2011 IEEE Symposium on Computers & Informatics, Kuala Lumpur, Malaysia, pp. 20–22 (2011)
3. Rick, J.: Performance optimizations of virtual keyboards for stroke-based text entry on a touch-based tabletop. In: Proceedings of the 23nd Annual ACM Symposium on User Interface Software and Technology (UIST 2010), pp. 77–86. ACM, New York (2010)
4. Mankoff, J., Abowd, G.D.: CIRRIN: A word-level unistroke keyboard for pen input. In: Proceedings of the 11th Annual ACM Symposium on User Interface Software and Technology, pp. 213–214 (1998)
5. Cechanowicz, J., Dawson, S., Victor, M., Subramanian, S.: Stylus based text input using expanding CIRRIN. In: Proceedings of the Working Conference on Advanced Visual Interfaces (AVI 2006), pp. 163–166. ACM, New York (2006)
6. Kristensson, P.O., Zhai, S.: Command Strokes with and without Preview: Using Pen Gestures on Keyboard for Command Selection. In: Proceedings CHI 2007: ACM Conference on Human Factors in Computing Systems, San Jose, California (2007)
7. Wobbrock, J.O., Myers, B.A., Kembel, J.A.: EdgeWrite: A stylus-based text entry method designed for high accuracy and stability of motion. In: Proceedings of the ACM Symposium on User Interface Software and Technology (UIST 2003), Vancouver, British Columbia, pp. 61–70. ACM, New York (2003)
8. Yatani, K., Truong, K.N.: An Evaluation of Stylus-Based Text Entry Methods on Handheld Devices in Stationary & Mobile Settings. In: Proceedings of MobileHCI 2007: The 9th International Conference on Human-Computer Interaction with Mobile Devices and Services, Singapore, pp. 145–152 (2007)
9. Perlin, K.: Quikwriting: continuous stylus-based text entry. In: UIST 1998: Proceedings of the 11th Annual ACM Symposium on User Interface Software and Technology, pp. 215–216 (1998)
10. Isokoski, P., Raisamo, R.: Quikwriting as a multi-device text entry method. In: Proceedings of the Third Nordic Conference on Human-Computer Interaction (NordiCHI 2004), pp. 105–108. ACM, New York (2004)
11. Ward, D.J., Blackwell, A.F., MacKay, D.J.C.: Dasher—a data entry interface using continuous gestures and language models. In: Proceedings of the 13th Annual ACM UIST, pp. 129–137

Study on Haptic Interaction with Digital Map on Mobile Device

Daiji Kobayashi[1], Yoshitaka Asami[2], and Sakae Yamamoto[2]

[1] Department of Global system Design, Chitose Institute of Science and Technology, Bibi 758-65, Chitose Hokkaido, 066-0012 Japan
d-kobaya@photon.chitose.ac.jp
[2] Department of Management Science, Tokyo University of Science, Kagurazaka 1-3, Shinjuku-ku Tokyo, 162-8601 Japan
{y-asami,sakae}@ms.kagu.tus.ac.jp

Abstract. Touch panel operation is low-skilled interface for young individuals, thus the map on PDA (digital map) is used by the young individuals. However, we have not seen the less Japanese elderly or older persons use the digital map. Although our previous study pointed that Japanese elderly have an allergy to using PC, we conducted the simulation experiment and investigated the usability of our custom digital map for Japanese elderly. Although almost of the older participants could not get to the goal by using the digital map, we assumed that the haptic interaction supported planning route and understanding present location but the haptic interaction. Therefore, we observed the older participant's haptic interaction with paper map and digital map while they move to a goal position in residential area. From the results of the observational study, the relation between the haptic interaction and their performance was revealed.

Keywords: haptic interaction, digital map, touch panel operation, navigation.

1 Introduction

1.1 Background

Recently, personal digital assistance (PDA) with touch panel is in widespread use in Japan. The touch panel operation is low-skilled interface for young individuals, thus the map on PDA (digital map) such as Google Map is used by the young individuals. More, it is known that the mobile device with touch panel such as the iPad® are easy to use for elderly. Although our previous study pointed that Japanese elderly have an allergy to using PC, we conducted a simulation experiment and investigated the usability of our custom digital map on PDA for Japanese elderly. As the simulation results, almost of the older participants could not get to the goal using the digital map and we observed unintentional or intentional manipulation at different situation. However, the effectiveness of HIM such as tracing a route on a map by a finger could not be cleared [1].

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 443–449, 2011.

1.2 Classification of Haptic Interaction

Recently, haptic interaction by hands was classified based on physiological aspect such as surface-electromyogram signals and articular angles (e.g. Ishikawa et al. [2], 2010), and aimed to apply to robotics engineering, remote-control interface, virtual interface and so on. On the other hand, Gibson (1966) pointed that the relation between haptic perception and body movement [3]. Lederman et al. (1996) observed was classified the haptic perception from the proprioceptive aspect [4]. The haptic perception is moving fingers, touching, holding, and manipulating for recognizing the object.

1.3 The Concept of Haptic Interaction with Paper Map

When an individual move to the goal using a paper map, he/she accesses the map and get information for the mobility; for example, an operator could rotate the map when the operator knows or decides the direction to move. These kinds of action are for getting visual information easily from the map and the environment. On the other hand, the operator's HIM is such as tracing the path again and again by a finger on the map for confirming the route. As regards the aim of the haptic interaction, we assume that the movement of the finger promotes more imaging the shape of the path toward the goal than imaging by just visual information. Therefore, we think that the HIM, a kind of innate behavior, draw a distinction between the path the operator should progress and the other path in the operator's mental map. That is, the HIM supports the recognizing the path toward the goal which is the object the operator should recognize from the map.

1.4 Haptic Interaction with Digital Map

In recent years, the digital map on mobile device with touch panel such as Google Map on PDA, personal navigation device, and so on has been widely used. The interfaces of the digital map with touch panel look similar. For example, the digital map scrolls by flipping the digital map in the opposite direction of what the operator is supposed to. This way of manipulation is not just for digital map but for touch panel operation of general application software on mobile device with touch panel. Thus there is a possibility that the touch operation for manipulation of digital map makes the haptic interaction with paper map difficult. Therefore, the relation between HIM and the touch operation for digital map should be considered.

1.5 Aim of This Study

In general, Japanese elderly people who have an allergy to information technology tend to dislike the troublesome operation. Almost of touch panel operation designed based on the innate behavior. Thus the mobile devise with touch panel is easier to use for the elderly than PCs with keyboard and computer mouse. Therefore, the manipulation of digital map should be introduced the HIM if the HIM is effective for recognizing and memorizing the route in the map. Accordingly, the characteristics of the HIM and the relation to mobility is discussed in this paper.

2 Method

2.1 Overview

To reveal the the relation between the HIM and the mobility, we conducted simulation experiment in real environment and observed the performance of the participants using video cameras. The respective participants walked the same starting point to the same goal one at a time using a map. The map the participant had was a paper map or a digital map. The digital map was made based on the paper map and displayed on a PDA with touch panel. The manipulations by the participants were classified into some categories and considered the relation to their mobility.

2.2 Experimental Area

The experimental area was a residential area in Tokyo and has 400 × 300 square meters with crowded city blocks of wooden dwellings and temples.

2.3 Participant

The participants were 10 elderly individuals who were residents near the experimental area and respective participants were physically unimpaired. They were randomized into 2 groups (Group-A and Group-B). Group-A, 3 males and 2 females, has ages ranging from 61 to 76 (mean = 67.7, S.D. = 5.5). Group-B, 2 males and 3 females, has ages ranging from 66 to 81 (mean = 74.3, S.D. = 6.2). All participants were not PDA users.

2.4 Maps

In experiment, a paper map and a digital map was used. The paper map, A3 size sheets was made based on blank map published by Ministry of Land, Infrastructure, Transport and Tourism of Japan. The content of digital map was the same as the paper map and the devise for indicating the digital map was a PDA (NTT DoCoMo T-01A). The PDA has a 4.1 inches wide VGA display with touch panel and was running Microsoft Windows Mobile 6.1 Professional Japanese edition. The custom software for indicating the digital map on the PDA was developed using Adobe Flash. The digital map displayed on a PDA was slide-able upward, downward, left or right by rubbing a finger across the screen as well as the way for the common digital maps. In addition, the digital map included two semi-transparent buttons on the map for zooming in and zooming out. Both of the maps indicated 12 impassable points because of the following two reasons: one was preventing them from proceed to uncharted area; and the other was because this experiment was not for getting to the goal easily by the participants. Therefore, the participants had to concern about avoiding passing the impassable point using the map.

2.5 Procedure

The respective participants of Group-A tried to walk promptly using the paper map one at a time and the participants of Group-B promptly walked using the digital map as well as the Group-A. Before the trial, we instructed the participant to reach the goal within 30 minutes and indicated the starting point on the map. Further, the participant of Group-B was allowed to practice in an attempt to use the digital map before they started to the goal. The participant's manipulation of the map was observed by two experimenters. The experimenters followed the participant and recorded from a different viewpoint using video cameras (SONY Handycam HDR-SR12). When a participant inadvertently passes the impassable point, the experimenter diverted the participant without telling where the participant was. After a participant accomplished the trial, they answered a questionnaire we had prepared for the Group-A. The questionnaire consisted of 9 question items, which requested feedback about the difficulty of the trial and the interaction between the participant and the map. Meanwhile we had prepared another questionnaire for the Group-B, which requested feedback about the difficulty of trial and the usability of the digital map on PDA.

3 Results

As the results of experiment, one male and one female participant of each group could not get to the goal within 30 minutes. According to this result, 4 participants gave-up trial or could not get to the goal; therefore, each group included 3 successful individuals and 2 failures. Although the respective groups include failures, we observed all participants' behavioral data in order to extract the behavior relating to the interaction with the map.

3.1 Classification of Manipulation

The participants' behavior observed was classified into 6 types. The appearance frequencies (behaviors per minute) of all behavior are shown by Table 1. In this regard, the 'Zooming' and the 'Sliding Map' manipulation were for operating the digital map and the 'Folding' and the 'Spreading' manipulation were for operating the paper map. Further, the manipulation 'Steadying a Point' means applying thumb pressure to a point on the map where the participant wanted to memory for a while. The point was such as present location, the next destination, and so on. The 'Rotating' manipulation is such as facing the paper map or the PDA displaying the digital map in the direction the participant going. The aim of the 'Rotating' was matching the map and the real environment where the participant was located in order to orient.

Table 1 indicates that the frequencies of the 'Rotating' and the 'Steadying a Point' manipulation in Group-B were higher than Group-A. However, 'Tracing a Route' manipulation, was the only kind of HIM, was commonly observed in two groups. These 3 types of manipulation were observed in both groups; therefore we compared these frequencies between successful participants' and failures'. In this regards, there were too few participants to compare statistically and all participants' data is represented as the following figures.

3.2 The Manipulation Was Intentional?

The feedback from the Group-A showed a few manipulations was intentional action. According to the feedback, the 4 of 5 participants of the Group-A intentionally rotated the paper map in order to face the map in the direction the participant was walking. However, the other one participant of Group-A never rotate the map could not get to the goal in the trial. Meanwhile, the 'Steadying a Point' manipulation and the 'Tracing a Route' manipulation was observed in all participants' trials; however, one or two participants who could not reach the goal, reported that they could not remember these 2 types of manipulations. Therefore, in the case of these manipulation types, the participants who intentionally interact with the map succeeded the trial.

Table 1. Types of manipulations observed from the behavior of respective groups and the appearance frequencies. The number shown in parentheses indicates the standard deviation.

Manipulation	Group A	Group B
Rotating	0.975 (0.710)	1.493 (0.459)
Folding	0.049 (0.305)	N/A
Spreading	0.026 (0.214)	N/A
Steadying a Point	0.539 (0.372)	1.376 (0.950)
Tracing a Route	1.337 (0.817)	1.079 (0.593)
Zooming	N/A	0.391 (0.165)
Sliding Map	N/A	0.731 (0.623)

3.2 Usability of the Digital Map

The usability evaluation by the Group-B indicates that our custom digital map was usable and easy to use; however, a participant reported that touch panel interface was easy to use but the digital map was slightly difficult to use. The reason why such the digital map was usable, the all participants pointed, was to be able to magnify the map. Although the all participants were elderly or older individuals, they answered that touch panel was preferable alternative of human-computer interface. Therefore, the following results should be taken into account that it was not hard to try to use the digital map for the participants of Group-B.

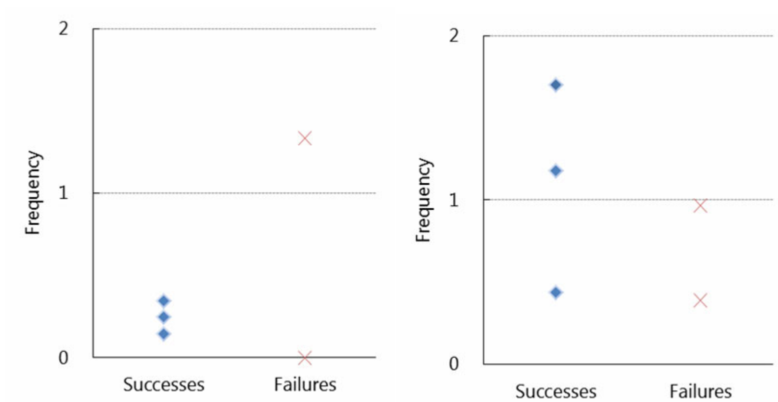


Fig. 1. Appearance frequency of the 'Rotating' manipulation in Group-A (left) and Group-B (right)

3.4 Frequency of ‘Rotating’ Manipulation

Figure 1 represents the frequencies of the ‘Rotating’ manipulation. The frequency of successes in Group-A is lower than successes’ in Group-B. These frequencies suggest that the ‘Rotating’ manipulation have less to do with the mobility and the type of the map.

3.5 Frequency of ‘Steadying a Point’ Manipulation

Figure 2 indicates that there is no relation between the maps or task success. Thus, it is suggested that the manipulation ‘Steadying a Point’ have less to do with the mobility and the type of the map.

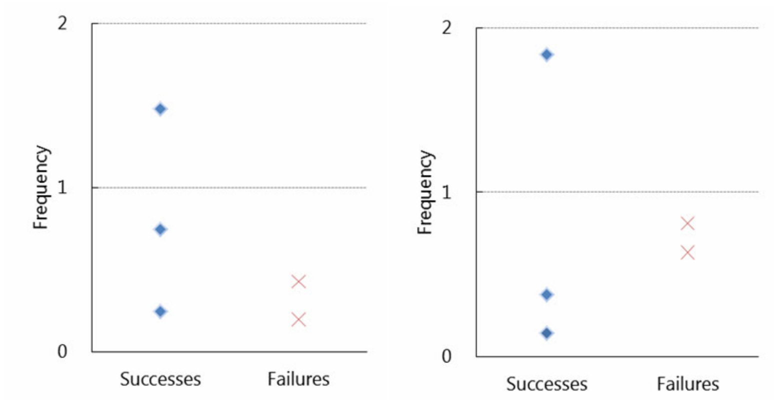


Fig. 2. Appearance frequency of the ‘Steadying a Point’ manipulation in Group-A (left) and Group-B (right)

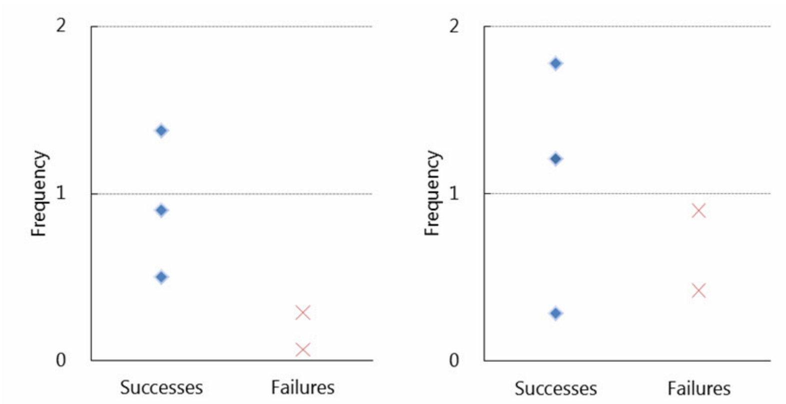


Fig. 3. Appearance frequency of the ‘Tracing a Route’ manipulation in Group-A (left) and Group-B (right)

3.6 Frequency of 'Tracing a Route' Manipulation

As Table 1 shows the frequency of the 'Tracing a Route' manipulation, the kind of HIM, observed with highest frequently in Group-A; however, the manipulation was the third highest frequently in Group-B. The differences of the frequency between successive participants and failures in each group are shown as Figure 3. Because of insufficient data, it is not certain but Figure 3 shows that the frequency of successes tends to be higher than failures' in both groups.

4 Discussion

The feedbacks from the participants suggest that intentional manipulation for the paper map leaded to successful mobility. However, the results also suggest that never the all manipulation but the 'Tracing a Route' intentional manipulation could relate to the successful mobility in spite of a few numbers of participants' data. Moreover, the results of the 'Tracing a Route' manipulation for digital map were as well as the results of using the paper map. Therefore, we assumed that such the 'Tracing a Route' manipulation could be an effective intentional manipulation and could be a kind of HIM.

The participants using the paper map made an effort to walk to the goal as much as possible, and they seemed to perform by decision-making process for mobility. The manipulation for the paper map or the digital map we observed could be such the intentional behavior in accordance with the decision-making or the strategy for mobility. In other words, the individuals who own the significant way for grasping the direction to move and the present location could reach the goal. If the 'Tracing a Route' manipulation led the navigator to smart mobility, the next digital map operation should include the HIM.

5 Conclusion

According to this study, we believe that the 'Tracing a Route' manipulation was a HIM. Further, it is possible that the HIM leads an individual to successful mobility. However, further study should be done in order to explain the characteristics of HIM more clearly.

References

1. Kobayashi, D., Yamamoto, S.: Cognitive Process of Evacuee with Map on Personal Digital Assistance. In: Lin, D.M., Chen, H. (eds.) Ergonomics for All, pp. 153–158. CRC Press, Boca Raton (2010)
2. Ishikawa, K., Toda, M., Sakurazawa, S.: Akita. J., Kondo, K.: Finger motion classification using surface electromyogram signals. In: 9th IEEE/ACIS International Conference on Computer and Information Science, pp. 37–42. IEEE Press, New York (2010)
3. Gibson, J.J.: he senses considered as perceptual systems. Boston, Houghton Mifflin (1966)
4. Lederman, S.J., Klatzky, R.L.: Haptic classification of common objects: Knowledge-driven exploration. Cognitive Psychology 22, 421–459 (1990)

Characteristics of Information Transmission Rates Using Noncontact Tactile Display

Kentaro Kotani, Masayoshi Hayashi, Nobuki Kido, and Takafumi Asao

Department of Mechanical Engineering, Kansai University
Kotani@kansai-u.ac.jp

Abstract. Recently, tactile interfaces, which are user interfaces that employ tactile perception for input/output, have been the focus of an increasing number of studies. However, tactile interface design guidelines to help optimize the characteristics of human tactile perception have not yet been fully developed. Thus, the objective of the study was to identify the influence of perception of phantom sensation on the location of stimulus presentation on the palm. For this purpose, a noncontact tactile display with an independent driven 12 by 12 matrix air-jet driver was constructed. The subjects orally responded one of following three perception levels, i.e., (1) a crisp phantom sensation was obtained, (2) incomplete phantom sensation was perceived due to scattered stimuli, and (3) only separate stimuli were perceived and no integrated perception was generated. Phantom sensation was perceived apparently at the proximal side of the index finger compared to the wrist areas. The obtained topography drastically changed between subjects, suggesting that each individual has different distribution of perceived phantom sensations. Future study included the range of securely perceived phantom sensations for determining the specifications of tactile displays.

1 Introduction

Tactile interfaces, which are user interfaces that employ tactile perception for input/output, have been the focus of an increasing number of studies. However, tactile interface design guidelines to help optimize the characteristics of human tactile perception have not yet been fully developed. It was indicated that the fundamental characteristics of tactile interface design are not properly understood [1]. Especially, we have been studied the characteristics of phantom sensation for tactile display, where a total of eight different shapers were identified with a less than 10 percents of error rates [2]. The individual differences varied although the same location of the palm was carefully arranged for the appropriate tactile stimulus presentation. It should be good to know the distribution of accurate perception of tactile stimuli by individuals which may offer quantitative specification for the design of tactile displays. Thus, the objective of the study was to identify the influence of perception of phantom sensation on the location of stimulus presentation on the palm.

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 450–453, 2011.

2 Methods

A noncontact tactile display with an independent driven 12 by 12 matrix air-jet driver was constructed as shown in Figure 1. Pressurized air was generated by an air compressor and delivered using solenoid valves. The air pressure was then regulated using a series of electro-pneumatic regulators. The intensity of the air-jet stimuli was controlled at 100kPa, and stimuli were applied to the center of the left palm. A total of six subjects consented to participate in the study. All of them were recruited from the university community and were right-handed. The tested area consisted of 81 locations of the palm (see Figure 2), each of which was generated by using phantom sensations consisted of stimuli generated by three independent nozzles. There were five trials for each location. The subjects orally responded one of following three

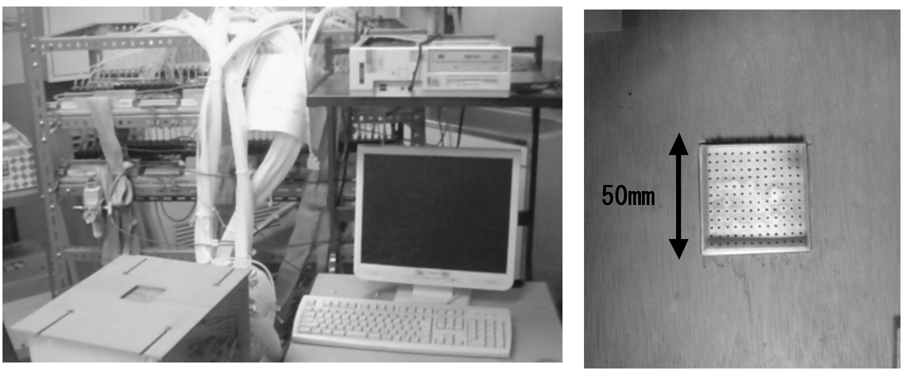


Fig. 1. Contactless tactile display (Left: system apparatus, Right: the display part where independent 144 air-jet nozzles were attached under the 50mm by 50mm-metal plate)

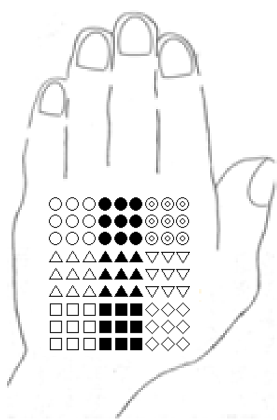


Fig. 2. Location of stimuli presentation given by phantom sensation. A total of 81 locations were chosen. The top middle point was aligned at the center point between the MP joint of the index finger and the MP joint of the little finger. Nine different symbolic patterns denote the regions tested as a location effect for statistical analysis.

perception levels, i.e., (1) a crisp phantom sensation was obtained, (2) incomplete phantom sensation was perceived due to scattered stimuli, and (3) only separate stimuli were perceived and no integrated perception was generated. Topography of perception of phantom sensation was generated for the individual to identify the tendency of the distribution of clearness of phantom sensation. The topography consisted of 255 levels of gradations ranging from white (no PS) and black (crisp PS).

3 Results

A two-way ANOVA was conducted to see the effects of subjects and the locations on the perception of tactile phantom sensations. The results revealed that significant individual differences $F(8, 432) = 5.810$, $p<.01$), and location differences ($F(5, 432) = 28.801$ $p<.01$) were found. Phantom sensation was perceived apparently at the proximal side of the index finger compared to the wrist areas, as shown in figure 3. The obtained topography drastically changed between subjects, suggesting that each individual has different distribution of perceived phantom sensations (see figure 3).

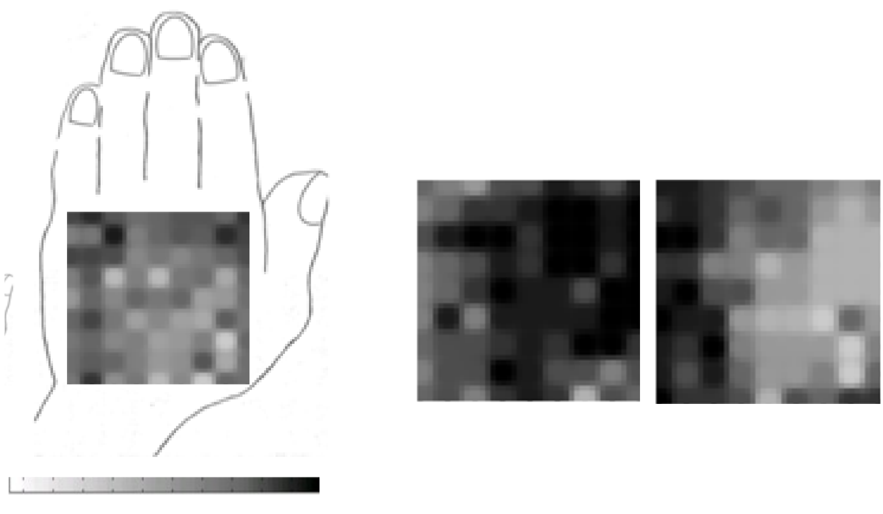


Fig. 3. Left: Grand average of topography of perceived clearness of phantom sensation (Grand average). In the figure legend, "No PS" means that a crisp phantom sensation was obtained, and "Clear PS" means that only separate stimuli were perceived and no integrated perception was generated. The clearly perceived locations for phantom sensation was distributed at the proximal side of the index and little fingers, whereas the center to the proximal side of the palm relatively shows that the subject perceived blur phantom sensation.Right: two topographies showing individual differences. Topography obtained from subject C showed clear PS over the palm area, whereas topography obtained from subject F showed a little PS on the palm and clear PS on the left of the palm (little finger side).

4 Discussion

Shimawaki, et al. [3] reported the detailed results associated with tactile two-point discrimination threshold by using calipers on the palm. Their results showed that distal part of the palm showed the highest sensitivity whereas the center of the palm had the lowest sensitivity. Our results can be interpreted that the air-jet stimuli applied at the distal part of the fingers relatively generated crisp phantom sensations, however, the individual differences for phantom sensation was so diverse that the tendency was rather marginal. Clear phantom sensations may be obtained under certain conditions including pressure levels and the shape of the palm as well as individual discrimination thresholds. The results implied that tactile transmission rates for noncontact tactile display heavily depended upon individuals thus the information using this device has been limited in current structure unless the information was presented at the outer edge of the palm close to the little finger; the users for this device may perceive information differently, compared with system as an alternate medium for visual information such as Braille. Future study included the range of securely perceived phantom sensations for determining the specifications of tactile displays.

Acknowledgements. This research was financially supported in part by the Kansai University Research Grants: Grant-in-Aid for Cooperated Work, 2010.

References

1. Suzuki, Y., Kobayashi, M.: Air jet driven force feedback in virtual reality. IEEE Computer Graphics and Applications 25(1), 44–47 (2005)
2. Hayashi, M., Kotani, K., Asao, T., Horii, K.: Recognition mechanism for tactile shape patterns presented by an air-jet driven tactile display. Annual Meeting of Japanese Ergonomics Society 45, 326–327 (2009) (Japanese)
3. Shimawaki, S., Sakai, N.: Comparison of blind and sighted subjects' volar tactile sensation using static and moving two-point discrimination tests. Japanese Journal of Mechanical Engineering, Part C 72(715), 147–152 (2005) (Abstract in English)

Multimodal Threat Cueing in Simulated Combat Vehicle with Tactile Information Switching between Threat and Waypoint Indication

Patrik Lif, Per-Anders Oskarsson, Björn Lindahl, Johan Hedström, and Jonathan Svensson

Swedish Defence Research Agency, (FOI) P.O. Box 1165, SE-581 11 Linköping, Sweden
{patrik.lif,per-anders.oskarsson,bjorn.lindahl, johan.hedstrom,jonathan.svensson}@foi.se

Abstract. We investigated four display configurations for threat cueing in a simulated combat vehicle. The display configurations were a tactile belt only; the tactile belt combined with 3D audio; two visual displays combined with 3D audio; and a multimodal configuration (the visual displays, the tactile belt and 3D audio combined). The tactile display was also used for navigation information. The participant's main task was to drive according to the navigation information, and when threat cueing onsets occurred, as fast as possible align the heading of the combat vehicle with the displayed direction of the threat. The tactile display thus switched between navigation and treat cueing information. Performance was overall best with the multimodal display. Threat localization error was smallest with the visual and multimodal displays. The response time was somewhat longer with the tactile belt only, and especially in the front sector. This indicates interference between the two tasks, when threat cueing onsets occurred at the same position as the navigation information. This should however not be a problem in a real combat vehicle, since the sound alert will most likely not be excluded. Thus, if coded correctly tactile information may be presented for both navigation and threat cueing.

1 Background

Human perceptual experience normally involves stimulation of several senses or perceptual systems and multisensory information that is complementary or redundant can among other things improve detection and alerting effects [e.g. 1]. In virtual and augmented reality applications, the experience of immersion can be enhanced by combining information introduced to the visual, auditive and tactile senses. Combined information in bimodal and multimodal displays is considered and used for multitasking and attention management [2, 3].

We know from previous studies that multimodal information can improve operator performance [4, 5, 6]. In [4] the driver of a simulated combat vehicle efficiently used multimodal threat cueing by combined visual head down (HDD), tactile and 3D audio displays. Whereas the use of the 3D audio only display led to some front-back confusions, these were neutralized by the addition of the tactile display. The response time

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 454–461, 2011.

with the HDD improved when it was combined with the tactile and 3D audio displays into a multimodal display, and the precision of the tactile display improved when it was combined with the HDD and 3D audio into a multimodal display. According to the results in [4], the 3D audio display did not seem to contribute much to these improvement, thus these improvements were mainly due to the combination of the HDD and tactile displays. When the visual display was changed from a HDD to a HUD the response time was comparable with the tactile display [5]. Furthermore, in both studies the overall perception of threat position was rated highest with the multimodal display [4, 5].

Even though multimodal displays can improve operator performance, switching attention between two tasks using tactile stimulation in a combat vehicle is not fully investigated. Previous research has however shown the possibility of simultaneous use of several tactile signals, for instance, for soldier communication [e.g. 7, 8] and for in-vehicle information [e.g. 9]. We thus believe that tactile information in a combat vehicle can be simultaneously used both for threat cueing and navigation information.

The purpose of the present study was to investigate performance of visual, tactile and multimodal threat cueing combined with tactile presentation.

2 Method

2.1 Participants

Twelve naïve non-military subjects with normal sight and hearing participated in the experiment. Their mean age was 22.3 years (18 – 26 years).

2.2 Apparatus

The experiment was performed in a simulated combat vehicle, CV90. The simulator was built on a Moog 6DOF200E motion platform, and the simulation was delivered and controlled by an FOI custom-made simulation engine (HiFi Engine). The out of the window view was presented on a 37'' plasma screen.
Visual threat cueing was presented with both a HUD (Head-Up Display) and a HDD (Head-Down Display) (Fig. 1).

The HUD consisted of green dynamic arrows overlaid on the left or right side of the central simulated out of the window view. The dynamics of the arrows included a pulsing flow, with speed decreased by reduced angular distance to the threat. When the threat was within sight in the window, the arrows disappeared and a vertical line appeared for alignment to the threat. The HDD below the out of the window view showed a top view of the vehicle, with horizontal threat position indicated by a 15º wide wedge-shaped indicator. The HDD display was always oriented in the direction of the vehicle, thus the pointer rotated dynamically while turning the vehicle (Fig 1).

Tactile threat cueing and waypoint navigation was presented by a torso belt with twelve tactors (approximate frequency 120 Hz) positioned on clock positions (i.e. each tactor covered 30° of the horizontal dimension at 0°, 30°, 60°, 90° etc.). The tactors were individually activated, and only one tactor was activated simultaneously. Two easily distinguishable tactile patterns were used, for tactile threat cueing and navigation information respectively. The tactile pattern for the navigation task was

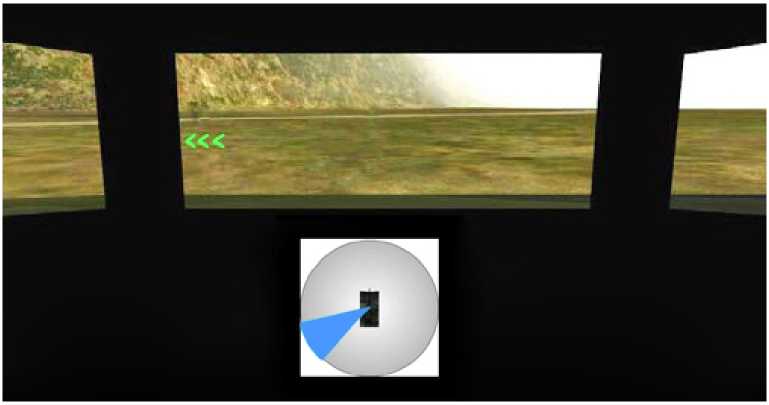


Fig. 1. Simulated out of the window view, with visual head-up display with arrow, indicating threat to the left and below head-down display showing threat position to the rear left of the own vehicle

continuously presented (on: 100 ms, off: 100 ms etc.) during driving until a threat appeared. The tactile pattern for threats was presented at threat cueing onsets (on: 90ms, off: 70ms, on: 100ms, off: 20ms, on: 100ms, off: 120ms etc.).

3D audio threat cueing was generated by an external soundcard and presented through headphones with a closed circumaural design. A LaserBIRD II headtracker was used to compensate for head movements in the 3D audio presentation.

Questionnaires with 7-point rating scales were used to assess mental workload; effort of receiving threat information; effort of receiving navigation information; if the navigation task interfered with threat handling; effort of performing the combined task of driving and handling the threat task; effort of driving; and perception of threat position in the different phases of the task (1 = very low, 7 = very high). The phases were at threat cueing onset (initial), when aligning the vehicle with the threat (middle), and when pushing the trigger button (final).

2.3 Experimental Design, Stimuli, and Procedure

The experiment had a within-subjects design with four display modes for threat cueing: visual+3D audio, tactile+3D audio, tactile-only, and multimodal (visual, tactile and 3D audio combined) and threat presentation in three sectors: front, back, and side. Waypoint navigation was given by the tactile display in all four display configurations (Table 1).

Table 1. Display configurations for waypoint and threat indication

Waypoint indication	Threat indication
Tactile	Visual + 3D audio
Tactile	Tactile + 3D audio
Tactile	Tactile-only
Tactile	Multimodal (visual+tactile+3D audio)

The participant's task was to maneuver the combat vehicle along a route, following waypoints indicated by a tactile pattern, and respond as quickly as possible when a threat was indicated. When the threat was perceived as being located straight ahead of the vehicle (indicated with a line in the visual display, vibrations in the front centre of the torso, or 3D audio sound positioned in the front of the vehicle) a trigger button was to be pushed. The threats were not visible, but textual feedback was given in the form of "hit" or "miss" overlaid on the simulated out of the window view. A hit was defined as vehicle heading aligned within ±10º from the threat direction. Each threat presentation was limited to the duration of 20 seconds or was terminated by the trigger button response. After the threat presentation was terminated the tactile navigation pattern was activated again. When the vibration pattern was on the side of the torso this indicated a new driving direction, and the participant should then turn the combat vehicle to get the vibration aligned with the centre of the torso. The simulated environment contained mainly open fields and some forests. The waypoints were placed in the terrain, thus it was not possible to follow roads.

The threats were presented in sectors relative to vehicle heading: at the front (0º ±40º but excluding the central ±10º), the sides (90º ±40º and 270º ±40º), and the back (180º ±40º). Direction of threat presentation was independent of vehicle orientation.

Localization error (LE) was defined as the absolute value in degrees of angular deviation of vehicle heading from threat direction when the trigger button was pushed. Response time (RT) was defined as elapsed time in ms from threat cueing onset to having turned the vehicle by 10°.

Each participant was exposed to four blocks of trials, one for each of the four display configurations. Each block consisted of 24 threat presentations, 8 randomly distributed in each of the 3 sectors (front, back, sides). The presentation order was balanced over participants.

First the background questionnaire was answered and a training session was performed with eight threat presentations for each display configuration. Then the four experimental blocks were performed, with one questionnaire after each block, and finally a summarizing questionnaire. The participants were told to turn the vehicle towards the displayed threat position and push the trigger button as fast as possible, and to perform the navigation task as accurately as possible.

3 Results

Performance data was analyzed with analysis of variance (ANOVA) with a 4 × 3 repeated measures factorial design (4 displays × 3 threat sectors). Subjective ratings of perception of threat position was analyzed with ANOVA with a 4 × 3 repeated measures factorial design (4 displays × 3 threat phases). The other subjective ratings were analyzed with ANOVA for a one-way repeated measures design (4 display configurations). Tukey HSD was used for all post-hoc testing.

Response Time (RT). The ANOVA showed a significant main effect of display configuration $F(3, 33) = 9.41$, $p < .0005$, and of sector, $F(2, 22) = 9.41$, $p < .005$, and a tendency of interaction effect of display configuration by sector, $F(6, 66) = 1.88$, $p = .097$ (Fig. 2).

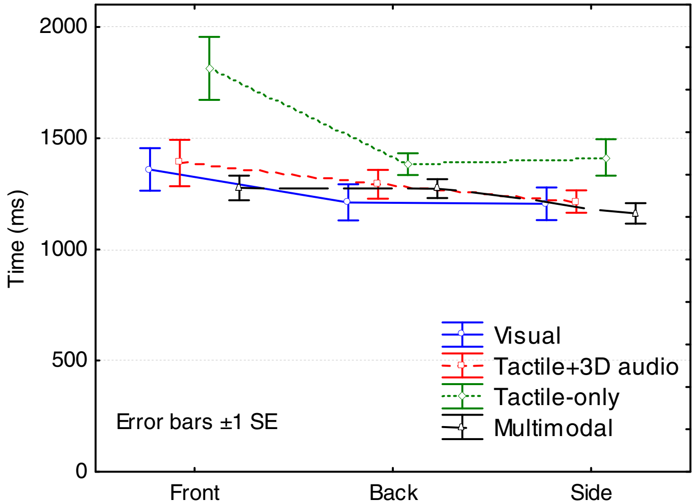


Fig. 2. Response time (RT) for the four display modalities for front, back, and side sector

Post hoc testing showed that the main effect of display was due to significantly longer RTs with tactile-only compared to the other display configurations (*p*s < .001). The main effect of sector was due to significantly longer RT in the front sector compared to the other sectors. The reason for the tendency of interaction effect was that the difference between the tactile-only and the other displays was largest in the front sector (*p*s < .01).

Localization error (LE). The ANOVA showed a significant main effect of display configuration $F(3, 33) = 63.2$, $p < .0001$, and a significant main effect of sector F(2, 22) = 3.6, $p < .05$.

Post hoc testing showed that the main effect of display was due to significantly larger LEs for both the tactile+3D audio (7.4 ±0.4) (Mean ±SE) and tactile-only (6.7 ±0.5) compared to both the visual+3D audio (2.3 ±0.3) and multimodal (2.4 ±0.3) displays. Post hoc testing also showed that the main effect of sector was due to small, but significantly higher LEs for threat cueing onsets presented in the side (5.1 ±0.4) compared to the back (4.4 ±0.2) ($p < .05$), but not compared to the front sector (4.6 ±0.2) ($p > .05$).

Perception of threat position. The ANOVA showed a significant main effect of display configuration, $F(3, 33) = 3.08$, $p < .05$, a significant main effect of threat phase, $F(2, 22) = 9.96$, $p < .001$, and a significant interaction effect of display configuration by threat phase, $F(6, 66) = 3.18$, $p < .01$ (Fig. 3).

Post hoc testing of the interaction effect showed that in the initial phase the ratings of the multimodal were significantly higher compared to the visual+3D audio display ($p < .05$) and a tendency of higher ratings compared to the tactile-only display ($p = .08$). In the middle phase the ratings of the multimodal display were only significantly higher compared to the tactile-only configuration ($p < .01$). In the final phase

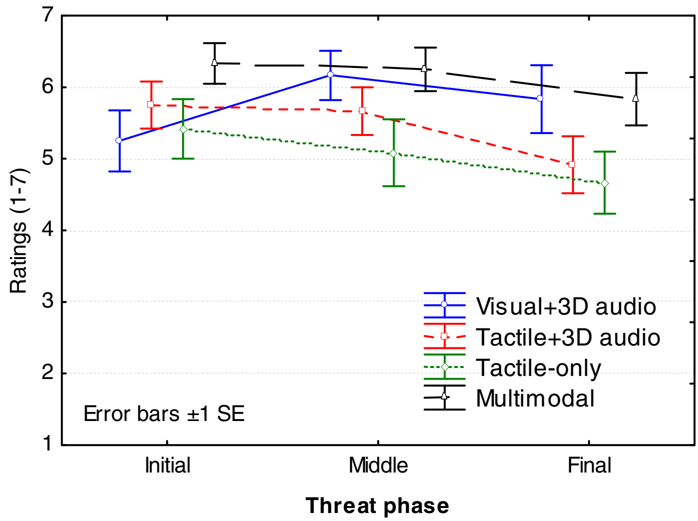


Fig. 3. Ratings of perceived threat position for the four display configuration at threat cueing onset (initial), when aligning the vehicle with the threat (middle), and when pushing the trigger button (final) (ratings made after simulation)

the ratings of the multimodal display were significantly higher compared to the tactile-only display ($p < .01$), and a tendency of higher ratings compared to the tactile+3D audio display ($p = .08$).

The main effect of display was due to overall higher ratings of the multimodal display compared to the tactile-only display ($p < .05$). The main effect of threat phase was due to overall lower ratings in the final threat phase.

Mental workload. The ANOVA showed no significant effect of display configuration, $F(3, 33) = 2.2$, $p > .05$). All means were on an intermediate level: visual (3.5 ±0.5) (Mean ±SE), tactile+3D (4.1 ±0.5), tactile-only (4.4 ±0.5), and multimodal (3.3 ±0.5).

Effort of threat information. The ANOVA showed no significant effect of display configuration, $F(3, 33) = 0.1$, $p > .05$). All means were low: visual (2.1 ±0.5), tactile+3D audio (2.0 ±0.3), tactile-only (1.9 ±0.4), and multimodal (2.2 ±0.4).

Effort of navigation information. The ANOVA showed a significant effect of display configuration, $F(3, 33) = 2.9$, $p < .05$). However, post hoc testing did not identify any significant differences between the displays, and all means were low: visual (1.1 ±0.1), tactile+3D audio (1.6 ±0.3), tactile-only (1.3 ±0.2), and multimodal (1.1 ±0.1).

Interference of navigation information. The ANOVA showed no significant effect of display configuration, $F(3, 33) = 1.8$, $p > .05$). All means were low: visual (2.3 ±0.4), tactile+3D audio (2.5 ±0.4), tactile-only (2.5 ±0.5), and multimodal (1.9 ±0.3).

Effort of combined tasks. The ANOVA showed no significant effect of display configuration, $F(3, 33) = 1.4$, $p > .05$). All means were rather low: visual (2.8 ±0.4), tactile+3D audio (3.2 ±0.5), tactile-only (2.7 ±0.3), and multimodal (2.6 ±0.4).

Effort of driving task. The ANOVA showed no significant effect of display configuration, $F(3, 33) = 1.1$, $p > .05$). All means were low: visual (1.5 ±SE), tactile+3D audio (2.2 ±0.4), tactile-only (1.8 ±0.3), and multimodal (1.6 ±0.2).

4 Discussion

The visual+3D audio threat cueing display configuration provided high precision and short response time. Since previous experiments [e.g. 4] have shown low precision with the 3D audio display, this indicates the importance of vision for precision tasks. Perceived threat position was however rated lower in the initial threat phase, during threat cueing onsets, compared to the multimodal display. This indicates the advantage of combining visual displays with tactile information for initial threat cueing.

The tactile threat cueing display configurations provided lower precision than the other displays. The tactile-only display provided longer RT than the other display configurations. That the tactile display+3D audio led to shorter RT, in level with the other display configurations, indicates the alerting value of the 3D audio sound. The longer RT with the tactile-only display in the front sector is probably due to interference with the tactile navigation task. The reason is most likely that at threat cueing onsets the participants were driving according to the navigation information, and thus received navigation information on the frontal tactors. Consequently, when the pattern changed to threat indication, there was a delay of about 400 ms before the participants noticed the change (compared to the other display configurations). According to the subjective ratings, the participants did not recognize this as a problem. This delay was remedied when tactile threat cueing was combined with 3D audio.

The multimodal threat cueing display provided high precision, short response time, and good perception of threat position in all three threat phases. That is, for all measures it was comparable, or better, than the best of the other displays. It is important that the rated mental workload of the multimodal display was only on an intermediate level, and not higher compared to the other displays.

The small but significant difference between the back and side sector for localization error was probably due to chance. A possibility is however that threat cueing onsets to the side of the vehicle triggered a more distinct initial turning, and thus a somewhat less careful final alignment with the threat.

According to the subjective ratings the participants did not find the presentation of navigation information as effortful, nor did they experience it as effortful to perform the combined task of navigating and handling threats, and they did not experience that the navigation information interfered with the threat handling. Threat cueing performance was at approximately the same level as in previous experiments without a tactile navigation task [4, 5]. Thus, the tactile navigation task did not seem to negatively affect threat cueing performance, except for a small impairment with the unimodal tactile-only display where threat cueing onsets could occur at the same position as the navigation information. Even if this impairment was only indicated by a tendency of an interaction effect, the post hoc test was significant. Although this effect was not formally identified by the analyses, we find it important to report this indication, since a delay of action in the order of 400 ms in a combat vehicle may be decisive. On the other hand, this problem is unlikely to occur in a real combat vehicle, since the sound alert for threat cueing will probably never be excluded.

The main conclusions of this study are, firstly, that the multimodal display with visual, tactile, and 3D audio displays combined can overcome display specific limitations and thus provide enhanced performance, which is in line with previous studies [4, 5]. Secondly, if coded correctly tactile information may be presented for both navigation and threat cueing.

References

1. Spence, C., Driver, J.: Crossmodal space and crossmodal attention. Oxford University Press, Oxford (2004)
2. Sarter, N.: Multiple-Resource theory as a basis for multimodal interface design: Success stories, quantifications, and research needs. In: Kramer, A.F., Wiegmann, D.A., Kirklik, A. (eds.) Attention: From Theory to Practice, pp. 187–195. Oxford University Press, New York (2007)
3. Wickens, C.D.: Multiple resources and performance prediction. Theoretical Issues in Ergonomics Science 3(2), 159–177 (2002)
4. Carlander, O., Eriksson, L., Oskarsson, P.-A.: Handling uni- and multimodal threat cueing with simultaneous radio calls in a combat vehicle setting. In: Stephanidis, C. (ed.) UAHCI 2007 (Part II). LNCS, vol. 4555, pp. 293–302. Springer, Heidelberg (2007)
5. Oskarsson, P.-A., Eriksson, L., Lif, P., Lindahl, B., Hedström, J.: Multimodal threat cueing in simulated combat vehicle. In: Proceedings of the Human Factors and Ergonomics Society 52th Annual Meeting, pp. 1287–1291. Human Factors and Ergonomics Society, Santa Monica (2008)
6. van Erp, J.B.F.: Tactile displays for navigation and orientation: perception and behaviour. PhD thesis Utrecht University, The Netherlands (2007) ISBN/EAN 978-90-393-4531-3
7. Allan, K., White, T., Jones, L., Merlo, J., Haas, E., Zets, G., et al.: Getting the buzz: What's new for tactile information delivery? In: Proceedings of the Human Factors and Ergonomics Society 54th Annual Meeting, pp. 1331–1334. Human Factors and Ergonomics Society, Santa Monica (2010)
8. Johnston, M., Hale, K., Axelsson, P.: Results from empirical testing of the system for tactile reception of advanced patterns (STRAP). In: Proceedings of the Human Factors and Ergonomics Society 54th Annual Meeting, San Francisco, CA, pp. 1335–1339. Human Factors and Ergonomics Society, Santa Monica (2010)
9. Van Erp, J.B.F., van Veen, H.A.H.C.: Vibrotactile in-vehicle navigation system. Transport Research Part F: Traffic Psychology and Behaviour 7(4-5), 247–256 (2004)

Design of Vibration Alert Interface Based on Tactile Adaptation Model to Vibration Stimulation

Yuki Mori, Takayuki Tanaka, and Shun'ichi Kaneko

Kita 14, Nishi 9, Kita-ku, Sapporo, Hokkaido, 060-0814, Japan
mori@ssc.ssi.ist.hokudai.ac.jp,
{ttanaka,kaneko}@ssi.ist.hokudai.ac.jp

Abstract. In this paper, we describe the influence of vibratory adaptation on vibration perception and its correction method. Vibration strength after change depends on the vibration frequency before change even if the vibration frequency itself does not change. We found that the higher the vibration frequency before change, the weaker the perceived vibration strength. To correct this perception gap, the frequency before change is fed back to the KI model that expresses the vibratory adaptation. We performed a simulation to show that output of the KI model is able to express changes in vibration caused by adaptation.

Keywords: Vibration, Adaptation, Sense of touch.

1 Introduction

Recent studies have focused on conveying information through the sense of touch, including vibrations, and tactile impressions. As an example, a system that conveys road-departure warnings and proximity warnings to drivers by steering wheel vibrations has already been put to practical use in cars.

Much information is displayed in a visual manner, but it is necessary to consider the user's line of sight and how disruptive it might be to, for example, driving car. In the case of information presented through audible sounds, the information is conveyed not only to the user, but also individuals located in the vicinity of the user, which might be undesirable for the given application. Information conveyed through vibrations does not have these drawbacks. Therefore, research has been conducted on a number of topics regarding the presentation of information through touch, such as [1][2][3].

In this paper, we present a vibration alert interface (VAI) that consists of multiple vibration frequencies. Conventional devices using vibration convey only one bit of information to indicate danger based on an on/off control of the vibration motor. The VAI aims to convey analog-like information by using changes in the vibration frequency; however, such changes in vibration are not always perceived due to human adaptation, which poses a problem for VAI applications. More specifically, even if the vibration frequency is the same, user perception changed based on vibratory adaptation.

In addition to this introductory section, this paper is organized as follows:

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 462–469, 2011.

Section 2 describes our experimental results regarding changes in vibration perception due to adaptation; Section 3 explains the KI model and our proposed correction method for the change in vibration perception induced by adaptation; in Section 4, we confirm through simulation that the KI model is able to express the vibratory adaptation and is effective in correcting the perception gap; and we conclude our paper and discuss future work in Section 5.

2 Change in Vibration Perception Induced by Adaptation

The VAI is a device that conveys information to users by using vibrations. Users recognize such information as proximity or danger from the VAI by perceiving the vibration strength, frequency, and pattern; however, a user's vibration perception changes due to disturbance [4][5][6]. Therefore, it is necessary to consider this disturbance in designing the VAI to keep the vibration perception of the user constant. For this, we conducted experiments to examine the change in vibration perception induced by vibratory adaptation in the form of disturbance.

2.1 Experimental Setup

For our experiments, we used magnitude estimation as an approach. Magnitude estimation is a scaling method used in psychophysics to evaluate a subject's perception of stimulus intensity. Our experimental method was as follows:

1. Subjects grasp the VAI vibrating at 110 Hz and are informed that the stimulus has a magnitude of 10.
2. The VAI is vibrated at frequencies 70, 90, 110, 130, or 150 Hz. The frequencies are randomly selected. Vibration at each of these frequencies lasts for 5 sec. These frequencies are labeled as frequency before change.
3. The frequency of VAI change. Subjects reported their perceived vibration strength relative to the standard strength of 10 after the change in frequency. These frequencies are labeled as frequency after change.

This experiment examined the change in vibration perception after the frequency change as a function of the vibration frequency before change. In this experiment, the vibration device shown in Fig. 1(a) was used as the VAI. The diameter and length of the vibration device are 4 cm and 10 cm, respectively; a vibration motor was mounted inside the device. The test involved 13 subjects, both males and females, with an average age of 24.5 years. Subjects wore eye masks and heard white noise with headphones during the experiment to ensure they were focused solely on their tactile senses.

2.2 Experimental Results

We examined the difference of vibration perception after change that is caused by the frequency before change by changing the frequency to each of the possible values noted above. The correlation coefficient between the frequency before change and the reported vibration perception after frequency change was calculated. On average, 5 correlation coefficients for 11 out of 13 subjects had a negative value. Moreover, all correlation coefficients for 10 of the 13 subjects were negative. A negative correlation coefficient means that the vibration perception became weak; this occurred when

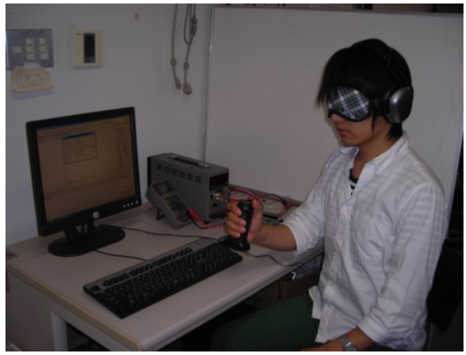


Fig. 1. Experimental setup and environment. (a) The VAI used in the experiment. (b) Experimental environment A vibration motor is mounted inside the vibration device.

the frequency after change was high, even if the vibration frequency after change was the same. On the other hand, a positive correlation coefficient means that the lower the frequency before change, the stronger is the vibration strength perceived by the subjects. The correlation coefficients and experimental results of all subjects are shown in Table 1 and Fig. 2.

Table 1. Correlation coefficients between perception after change and the frequency before change.

Frequency after change	Correlation coefficient
70	-0.93
90	-0.94
110	-0.99
130	-0.98
150	-0.97

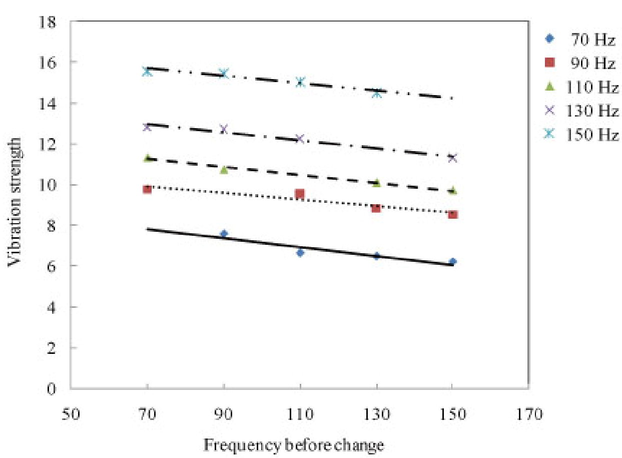


Fig. 2. Change in perceived vibration strength caused by vibration before change

Perceived vibration strength after change was induced by vibration before change, even if the vibratio frequency after change was the same. Subjectts perceived strength 10 when the frequency changed to 110 Hz, from 150 Hz, while subjects perceived strength 12 when the frequency changed to 110 Hz, from 70 Hz.

All correlation coefficients were close to 1.0; however, it was cleary shown that subjects perceived vibration s much more prominently when the frequency before change was low. In Fig. 2, the x-axis represents the frequency before change and the y-axis represents the perceived vibration strength after change. The slope of lines consistently point downward and this also shows how the perceived vibration strength changes with the frequency before the change.

3 KI Model

3.1 Configuration of Arrayed Neurons

In section 2 above, we described our experimental results in which the vibratory adaptation changed the perception of vibration. To convey correct information to a user, the perception gap due to adaptation has to be eliminated. Therefore, we propose incorporating a feedback mechanism for frequency before change into the vibratory adaptation model, as shown in Fig. 3. We chose the KI model as the adaptation model [7].

Fig. 4 shows a schematic configuration of arrayed neuron of the KI model. The KI model consists of excitatory and inhibitory neurons complement one another. An excitatory neuron generates a positive impulse and an inhibitory neuron generates negative impulse. Each receptor is connected with an excitatory and an inhibitory neuron. The two neurons are combined to form a 2nd-order neuron. When a signal input is applied to a receptor, both the neurons generate a nerve impulse and information is captured by the 2nd-order neuron.

3.2 Formulation

A single pulse generated by each of the excitatory and inhibitory neuron is given as

$$u_{\pm}(x,t) = \pm\frac{1}{2}Q_0\frac{R\alpha_{\pm}}{\sqrt{\pi t}}\exp\left(-\frac{\beta_{\pm}}{\alpha_{\pm}{}^2}t - \frac{\alpha_{\pm}{}^2x^2}{4t}\right) \tag{1}$$

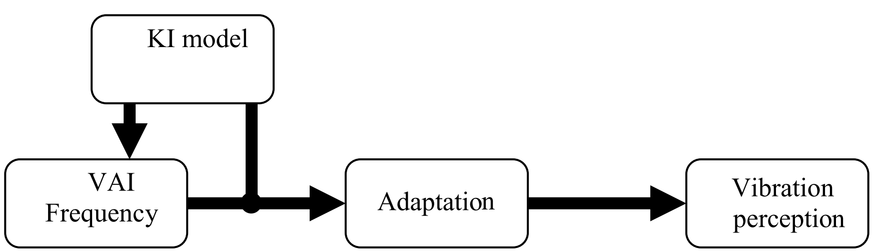


Fig. 3. Correcting the perception gap due to adaptation; the frequency before change is fed back and the VAI vibrates at the corrected frequency as calculated by the KI model

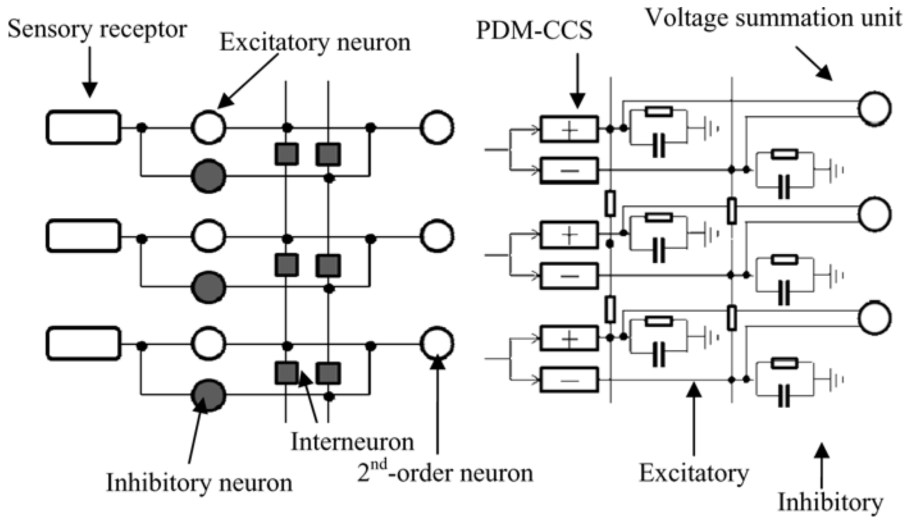


Fig. 4. Configuration of arrayed neurons for KI model and an equivalent circuit

where + and – correspond to the excitatory and inhibitory neurons, respectively, $Q_0/2$ shows pulsed current at x = 0, R is the resistance, α is the diffusion speed, and β/α^2 is the attenuation speed. These pulses are accumulated and given at the 2^{nd}-order neuron as

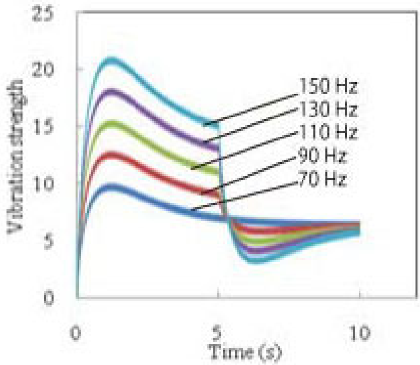
$$u(x,t)=\sum_{i=1}^{N}\left[u_+(x,t+iT_+ +\tau)+u_-(x,t+iT_-)\right] \tag{2}$$

where τ is the firing time difference between the excitatory and inhibitory neurons, and T is the pulse interval form the sensory receptor.

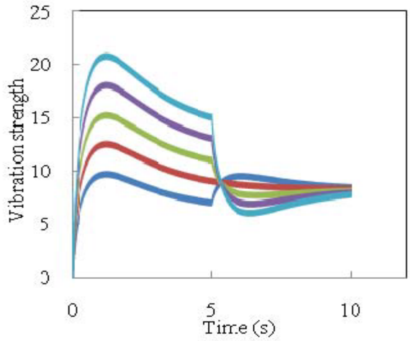
4 Simulation Using KI Model

To calculate the correction frequency using KI model, the KI model needs to express the change in vibration perception caused by vibration before change. If the KI model expresses the vibratory adaptation, the correction frequency is calculated using the frequency before change.

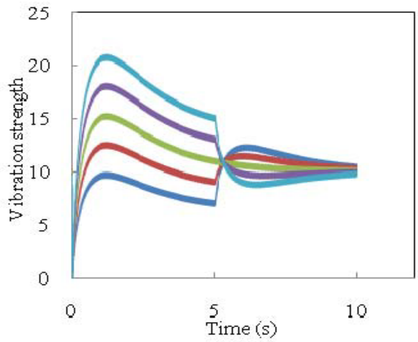
We conducted a simulation using KI model to examine the adaptation effect in this KI model. In this simulation, the vibration frequency was input into a sensory receptor, and the output was accumulated at the 2^{nd}-order neuron. This output is taken as the perceived vibration strength. The input frequency changed at the 5-second mark. Our simulation results are shown in Fig.5.



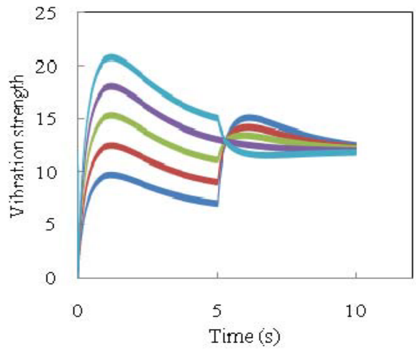
(a) 70 Hz



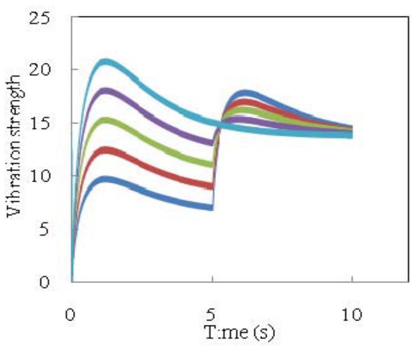
(b) 90 Hz



(c) 110



(d) 130



(e) 150

Fig. 5. Simulation using the KI model of the change in vibration strength due to the frequency before change. The vertical axis represents the perceived vibration strength after change. The frequency of 70-150 Hz changed to one frequency at the 5-second mark. The vibration strength after the change was changed due to the frequency before change, although the frequency after change was the same. Parameters: $\alpha_+ = 0.58, \alpha_- = 0.30,\ \beta_+ / \alpha_+^{\ 2} = 1.05,\ \beta_- / \alpha_-^{\ 2} = 0.52$.

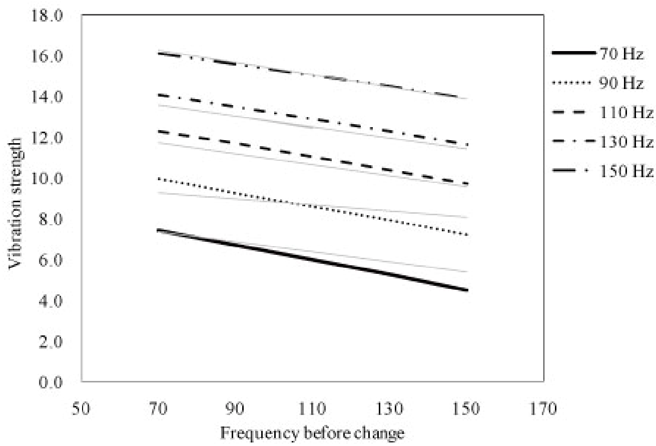


Fig. 6. A comparison of simulation results with experimental results; black lines represent simulation results, whereas gray lines represent experimental results from Fig. 2

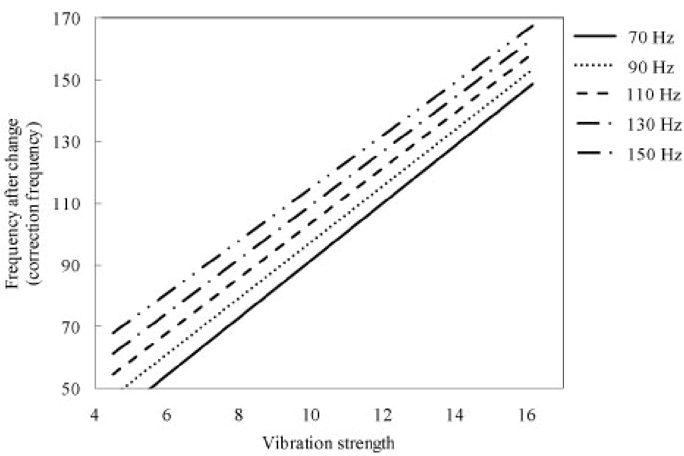


Fig. 7. Correction frequency calculated by the vibration strength and the frequency before change. When the frequency before change is 70 Hz and the VAI wants user to perceive the vibration strength 12, the correction frequency become 110 Hz.

Fig. 5(a) shows that the vibration perception for different values of frequency before change when the frequency after change was 70 Hz. Naturally, the vibration strength before change was different because the frequency before change was different; however, the vibration strength after change was different even though the frequency after change was the same. This is owing to the influence of adaptation, which was also confirmed for other frequencies, as shown in Fig. 5(b)-(e).

Fig. 6 shows a comparison of the simulation results with our experimental results. The slopes of the simulation results point downward similar to our experimental results. Further, the KI model seems to express the adaptation effect, because both sets of results are close to one another. From this simulation, we found that the KI model was able to express the change in vibration perception caused by adaptation.

Using the KI model, we can calculate the correction frequency to eliminate the change in vibration perception due to the adaptation effect because the KI model was able to express the change in vibration perception caused by adaptation. The vertical axis and the horizontal axis from Fig. 6 change to the frequency after change and vibration strength, respectively, we obtain Fig. 7. Fig. 7 shows the correction frequency calculated by the vibration strength and the frequency before change. Therefore, the KI model can calculate the correction frequency by input the vibration strength that the VAI wants to convey and the frequency before change.

5 Conclusion

In this paper, we examined the change in vibration perception by adaptation. We found that vibration perception after change was changed by the frequency before change when the vibration frequency changes. Moreover, owing to the adaptation effect, the higher the vibration frequency before change, the weaker the perceived vibration strength. To correct this effect of adaptation, we proposed a correction method that feeds back the frequency before change to the KI model. We trust that the KI model is able to successfully calculate the correction frequency, because through simulation the KI model successfully expressed the change in vibration perception caused by adaptation. In our future work, we ascertain the effectiveness of this feedback correction by using the KI model in correction experiments.

References

1. Yao, H.S., Grant, D., Cruz, M.: Perceived Vibration Strength in Mobile Devices: the Effect of Weight and Frequency. IEEE Transaction on Haptics (2009)
2. Morioka, M., Griffin, M.J.: Magnitude-dependence of equivalent comfort contours for fore-and-art, lateral and vertical hand-transmitted vibration. Journal of sound and vibration 295, 633–648 (2006)
3. Morioka, M., Griffin, M.J.: Frequency dependence of perceived intensity of steering wheel vibration: effect of grip force. In: Second Joint EuroHaptics Conference and Symposium on Haptic Interface for Virtual Environment and Teleoperator Systems, WHC 2007 (2007)
4. Nishiyama, K., Watanabe, S.: Temporary Threshold Shift of Vibratory Sensation After Clasping a Vibrating Handle. International Archives of Occupational and Environmental Health 49, 21–33 (1981)
5. Maeda, S., Griffin, M.J.: Temporary threshold shifts in finger tip vibratory sensation from hand-transmitted vibration and repetitive shock. British Journal of Industrial Medicine 50, 360–367 (1993)
6. Mori, Y., Tanaka, T., Kaneko, S.: Design of Vibration Alert Interface based on Cibration Strength Considering Skin Deformation with Respect to Grip Force. Human Interface Society 12(2), 103–111
7. Katagiri, Y., Aida, K.: Simulated nonlinear dynamics of laterally interactive arrayed neurons. In: Optomechatronic Technologies 2008 SPIE Proceedings (2008); Smith, T.F., Waterman, M.S.: Identification of Common Molecular Subsequences. J. Mol. Biol. 147, 195–197 (1981)

Applicability of Touch Sense Controllers Using Warm and Cold Sensations

Miwa Nakanishi[1] and Sakae Yamamoto[2]

[1] Keio University, Fac. of Science & Technology, Dept. of Administration Engineering, Hiyoshi 3-14-1, Kohoku, Yokohama 223-8522, Japan
miwa_nakanishi@ae.keio.ac.jp
[2] Tokyo University of Science, Fac. of Engineering, Dept. of Management Science, 1-3 Kagurazaka, Shinjuku, Tokyo 163-8522, Japan
sakae@ms.kagu.tus.ac.jp

Abstract. At present, designing controllers considering the physical aspect, such as a user's muscle load or range of movement, and the cognitive aspect, such as easy handling or simple functions, has become a basic requirement, and many products have satisfied these aspects. Thus, conversely describing, because it is difficult to find differences among products using these aspects for customers, manufacturers are seeking ways to effectively appeal to users in the psychological aspect. A touch sense is expected to produce new interaction that meets our psychological satisfaction beyond visual and auditory senses. In this study, aiming at producing the reality effectively in user–system interaction, we examine the applicability of controllers with warm and cold sensation that match users' images of a change in system conditions. In particular, 1) we examine if there is a relationship between the cold or warm sensation that users feel at their fingertips and an image of changes in system conditions, and 2) we examine if a slider-type controller with warm and cold sensations helps a user's finger to slide correctly and naturally. Furthermore, on the basis of the results of this study, we propose some concepts for applying a touch sense to actual situations.

Keywords: Touch sense, Warm and cold sensations, System conditions.

1 Introduction

In an overwhelming majority of user–system interaction, controllers that use hands as a modality to input information from a user to a system are adopted. Human hands are good at detail work and have a fine touch sense [1]-[4]. With advancing technologies, many controllers, from conventional handles and levers to the latest mice and touch pads, have appeared, and much ergonomic design has been applied to them. At present, designing controllers considering the physical aspect, such as a user's muscle load or range of movement, and the cognitive aspect, such as easy handling or simple functions, has become a basic requirement, and many products have satisfied these aspects. Thus, conversely describing, because it is difficult to find differences among

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 470–477, 2011.

products that use these aspects for customers, manufacturers are seeking ways to effectively appeal to users in the psychological aspect.

Regarding user–system interaction, it is well known that information is input from a system to a user via eyes or ears and that from a user to a system via hands or legs. However, Gibson [5][6] found that touching an object with hands could produce complex and rich information; hence, hands are a very useful modality that enable a user to obtain information from and provide information to a system simultaneously. Moreover, Norman [7] referred to the importance of embedding "reality" in designing a tool, and Katz [8] suggested that a touch sense with hands could give a more impressive reality to a user than the other senses. Therefore, a touch sense is expected to produce new interaction that meets our psychological satisfaction beyond visual and auditory senses.

In an experimental study to examine the possibility of using controllers with impressive reality, Nakanishi, et al. [9] found that users could recognize roughness or smoothness of the surface of a simple switch by just pushing it with their index finger, and that there were relationships that users shared between some of the images of system conditions and the roughness or smoothness that users felt when touching the surface of switches. (For example, users commonly thought of "start up" by the switch with a smooth surface and "stop" by that with a rough surface.) Furthermore, Nakanishi, et al. [10] carried out an experiment in which users experienced the remote controller of an air conditioner to which the above findings were applied, and obtained the results that attaching roughness and smoothness to the surface of individual buttons on the controller so that each matched the users' images made them feel more comfortable and relieved.

In this study, focusing on two additional viewpoints, we explore the application of a touch sense to controllers. The first point focuses on the warm and cold sensations felt at the fingertips and examines the applicability of the sensation to controllers. The second point focuses on active finger movement as applied to recent mobile phones, and in particular, examines finger slide in controlling system conditions. This study aims at producing the reality effectively in user–system interaction, and examines the applicability of controllers with warm and cold sensations that match users' images of a change in system conditions. In particular, 1) we examine if there is a relationship between the cold or warm sensation that users feel at their fingertips and an image of changes in system conditions and 2) we clarify if a slider-type controller with warm and cold sensations helps a user's finger to slide correctly and naturally.

2 First Experiment

In the first experiment, we tried to clarify whether there is a common image on a change in system conditions that occurred when users felt a warm or cold sensation at their fingertip while it was sliding on the controller.

2.1 Method

Experimental apparatus. We prepared an oblong panel composed of three pieces of thermo modules (Peltier devices) covered with an aluminum plate on which a user's finger slid (Fig. 1). A piece of thermo module uses the Peltier effect, which states that

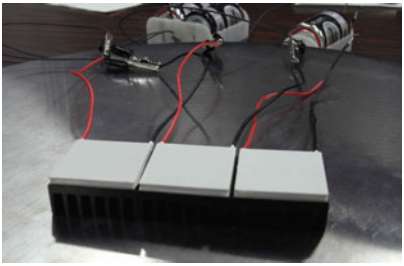


Fig. 1. Oblong panel composed of three pieces of thermo modules (Peltier devices)

heat transfers from one side to another when a current is made to flow at two junctures and that it is possible to absorb and radiate heat by changing the current direction. Each part of the oblong panel was designed to be warmer or cooler than the normal temperature by connecting each of the three modules to one or more dry batteries. The room temperature was adjusted to 25–26°C, so that the temperature of the panel surface was normally 25–26°C without passing a current.

In order that a subject sliding a finger on the oblong panel could feel a gradual change in temperature, different numbers of dry batteries were connected to each module. Table 1 shows six patterns of the temperature change. For example, when a subject slid a finger left to right on the oblong panel of pattern I-iii, he could feel a gradual temperature change, from cold to warm. The surface temperatures in Table 1 were measured using a surface thermometer (905-T2, manufactured by Testo).

Table 1. Six patterns of arrangement of thermo modules

			Surface temperature (°C)		
			Left	Center	Right
Increase in temperature	Pattern I-i	Low -> Normal	13	19	25
	Pattern I-ii	Normal -> High	26	37	48
	Pattern I-iii	Low -> High	19	27	37
Decrease in temperature	Pattern II-i	High -> Normal	48	36	25
	Pattern II-ii	High -> Normal	25	19	15
	Pattern II-iii	Low -> Normal	37	26	18

Experimental task. The subjects positioned their index finger on the right or left side of the panel and slid it toward the opposite side (Fig. 2). They performed the task using the six patterns shown in Table 1 in two manners: left to right and right to left. To eliminate the order effect, the 12 experimental conditions were tested randomly.

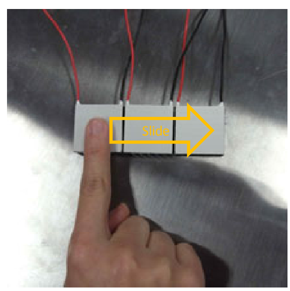


Fig. 2. Sliding the index finger on the thermo modules

After performing the tasks under each experimental condition, they answered questions about the relationship between the feeling of temperature change at their fingertip and the image of change in system conditions in the following manner. Twenty pairs of words representing system conditions, which were composed on the basis of our previous study [9] (Table 2), were presented to them, and they selected the side of a pair word that corresponded more to the temperature change they experienced at their fingertip. For example, when they slid their finger left to right on the panel of pattern I-iii and evaluated the experience on the image of a system condition of "Forward/Backward," if they felt that the experience matched with the image of "Backward," they selected it; however, if they felt that the experience matched with the image of "Forward," they selected it. While answering each pair word, they were allowed to repeatedly experience the temperature change in the panel under an experimental condition.

Table 2. Pair words representing system conditions

Forward/Backward	Near/Far
Up/Down	Stable/Instable
Large/Small	Sharp/Blunt
Fast/Slow	Light/Dark
Long/Short	Soft/Hard
High/Low	True/False
Heavy/Light	Play/Pause
Open/Close	Start/Stop
Right/Left	On/Off
Front/Back	Increase/Decrease

Subject. The subjects included 16 students aging 22 to 26. We obtained the informed consent of all participants.

2.2 Results

Fig. 3 shows the answers given for the pair words "Open/Close." The results indicate that most of the subjects thought of the image "Open" from the pattern of a temperature rise and the image "Close" from that of a temperature fall. However, we cannot find significant characteristics in the pattern where the temperature changed from high to normal or low to normal. Table 3 summarizes all results including the answers to the other pair words. In this table, the proportion that a left word was selected was represented with a negative value, and the proportion that a right word was selected was represented with a positive value. If the bias was symmetrical and the proportion was over 80%, an asterisk was placed in the right row. As this table shows, the images of system conditions represented by all pair words except "Stable/Instable" and "Sharp/Blunt" were related to the temperature change. Moreover, it was confirmed that for a temperature change from high to normal or low to normal, it was difficult to create an image.

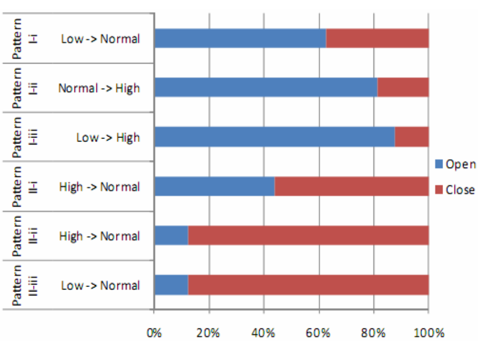


Fig. 3. Proportion of the answers given to the pair word "Open/Close"

Table 3. Proportion of the answers given to all pair words

		Forward/Backward		Up/Down		Large/Small		Fast/Slow		Long/Short		High/Low		Heavy/Light		Open/Close		Right/Left		Front/Back	
Pattern II-iii	Low -> Normal	88%	13%	100%	0%	81%	19%	63%	38%	69%	31%	100%	0%	50%	50%	81%	19%	88%	13%	63%	38%
Pattern II-ii	High -> Normal	56%	44%	75%	25%	44%	56%	56%	44%	50%	50%	56%	44%	25%	75%	63%	38%	56%	44%	56%	44%
Pattern II-i	High -> Normal	88%	13%	94%	6%	69%	31%	63%	38%	81%	19%	69%	31%	19%	81%	88%	13%	63%	38%	69%	31%
Pattern I-iii	Low -> High	31%	69%	6%	94%	19%	81%	44%	56%	19%	81%	13%	88%	56%	44%	13%	88%	38%	63%	13%	88%
Pattern I-ii	Normal -> High	50%	50%	50%	50%	63%	38%	44%	56%	38%	63%	56%	44%	75%	25%	44%	56%	56%	44%	56%	44%
Pattern I-i	Low -> Normal	25%	75%	0%	100%	6%	94%	38%	63%	31%	69%	19%	81%	50%	50%	13%	88%	25%	75%	25%	75%
Biased or not		*		*		*		*		*		*		*		*		*		*	

		Near/Far		Stable/Instable		Sharp/Blunt		Light/Dark		Soft/Hard		True/False		Play/Pause		Start/Stop		On/Off		Increase/Decrease	
Pattern II-iii	Low -> Normal	56%	44%	44%	56%	56%	44%	94%	6%	88%	13%	63%	38%	88%	13%	81%	19%	100%	0%	94%	6%
Pattern II-ii	High -> Normal	50%	50%	81%	19%	63%	38%	56%	44%	69%	31%	63%	38%	69%	31%	75%	25%	75%	25%	44%	56%
Pattern II-i	High -> Normal	75%	25%	56%	44%	56%	44%	81%	19%	94%	6%	75%	25%	100%	0%	88%	13%	94%	6%	94%	6%
Pattern I-iii	Low -> High	13%	88%	69%	31%	25%	75%	6%	94%	19%	81%	56%	44%	6%	94%	13%	88%	13%	88%	0%	100%
Pattern I-ii	Normal -> High	50%	50%	31%	69%	56%	44%	56%	44%	38%	63%	25%	75%	50%	50%	44%	56%	50%	50%	56%	44%
Pattern I-i	Low -> Normal	31%	69%	56%	44%	38%	63%	13%	88%	6%	94%	25%	75%	6%	94%	13%	88%	0%	100%	0%	100%
Biased or not		*						*		*		*		*		*		*		*	

3 Second Experiment

Next, we examined if the slider-type controller with warm and cold sensations that was designed on the basis of the results of the first experiment helped a user's correct and intuitive operation.

3.1 Method

Experimental apparatus. We prepared a mimic slide-type controller composed of five pieces of thermo modules (Peltier devices) covered with an aluminum plate. The number of dry batteries connected to each module and the direction of each current were fixed, as shown in Fig. 4, to set the temperature of each module to as shown in Fig. 5, so that a subject who slid a finger on the controller felt a temperature change (increase or decrease) from one side to another.

Experimental task. The subjects were given various words that indicated a change in system conditions (Table 4). They were asked to consider the given words as an

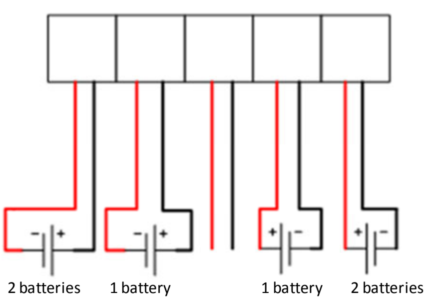


Fig. 4. Number of dry batteries connected to each module and the direction of each current

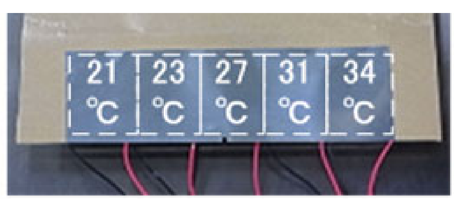


Fig. 5. Temperature of each module

Table 4. Words meaning change in system conditions

Up	A little up	A little down	Down
Increase	A little increase	A little decrease	Decrease
Open	A little Open	A little Close	Close
Light	A little light	A little dark	Dark
High	A little high	A little low	Low
Large	A little large	A little small	Small
Strong	A little strong	A little weak	Weak
Forward	A little forward	A little backward	Backward

operation purpose and to operate the slider-type controller. These words were composed of eight types of system conditions and four grades for each type, which were selected from the pair words used in the first experiment. The subjects positioned their index finger on the center of the slider-type controller and slid it to the position of a temperature that they thought matched the image of the given word; that is, the operation purpose. For example, when they were given the operation purpose "A little open," they slid their finger from the center toward the right or left, feeling warm or cold, and stopped at the position they thought matched the image of the operation purpose. The subjects performed the task 64 times, with each of the above operation purposes, using two types of controllers, in which the temperature increased from left to right or decreased from left to right. To eliminate the order effect, the operation purposes and the types of controllers were given to the subjects randomly. After each trial, we recorded the position where their finger stopped.

Subject: The subjects included 21 students aging 22 to 26. We obtained the informed consent of all participants.

3.2 Results

Fig. 6 shows the correlation between four grades of operation purposes concerning "Open/Close" and the temperature at which the subjects' finger stopped. From the results of the one-way ANOVA, we found a significant effect of the grades of the operation purposes on the temperature at which they stopped their finger ($F = 32.45$, $p < .05$). This means that for the operation purpose "Open," their finger intuitively moved toward the direction of high temperature and

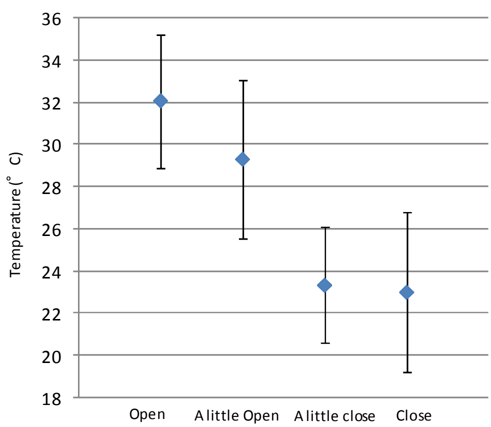


Fig. 6. Correlation between four grades of the operation purposes concerning "Open/Close" and the temperature at which the subjects' finger stopped

for the operation purpose “Close,” their finger intuitively moved toward the direction of low temperature. Only one subject showed the opposite reaction; thus, it was a common characteristic among the subjects. Moreover, the correlation coefficient 0.73 suggests that their finger stopped corresponding to the degree of the temperature. However, because the length of the finger movement was relative to the change in temperature in this experiment, it was understood that the correspondence of the degree between the system condition and the temperature described above was caused by both factors. Table 5 shows the correlation between the other operation purposes and the temperatures at which the subjects’ finger stopped. The results indicated that the subjects’ finger slid in common directions according to the increase or decrease in the temperature when they were given each operation purpose concerning the eight images of changing system conditions.

Table 5. Correlation between four grades of the operation purposes and the temperature at which the subjects’ finger stopped

				F-value	*<.05
Up	A little up	A little down	Down	77.55	*
Increase	A little increase	A little decrease	Decrease	67.13	*
Open	A little Open	A little Close	Close	32.45	*
Light	A little light	A little dark	Dark	42.74	*
High	A little high	A little low	Low	28.89	*
Large	A little large	A little small	Small	100.37	*
Strong	A little strong	A little weak	Weak	39.43	*
Forward	A little forward	A little backward	Backward	34.4	*

4 Application

This study’s results suggest the applicability of the controller using warm and cold sensations to the following situations as more impressive user interfaces. First, we can transfer the application to controllers in a car. Although the drivers have to pay attention to the outside environment for safety, it is also necessary to operate the air conditioner, the audio, and the navigation in order to enjoy driving. A touch sense can be effectively applied to user interfaces that are difficult to pay attention to while operating the vehicle. For example, if warm and cold sensations is applied to the slider-type interface to control the volume of the audio, it might provide drivers with a more intuitive and enjoyable experience in the car. Second, it will be effective to use a touch sense for user interfaces to operate some functions of machines in industrial fields. In general, touch panels are installed in some of the latest industrial machines for setting the amount of work or patterns of processes. However, in the industrial fields, the work environment is often much different from that in offices, in terms of lighting and noise. In such cases, if workers can control the amount of work or patterns of processes by using a touch panel with warm and cold sensations that match their intuition, human error can be reduced and relief and comfort can be enhanced.

5 Conclusion

In this study, aiming to effectively apply a touch sense to controllers in human–system interaction, we examined the usefulness of applying warm and cold sensations that matched users' intuition to a slider-type interface in order to control system conditions. The results suggested the following ideas. First, 18 types of images concerning system conditions were related to warm and cold sensations common to users. Second, when users operated a slider-type controller with operation purposes such as "Open/Close" and "Up/Down," a warm or cold sensation navigated their finger in the direction that matched the operation purpose.

Furthermore, on the basis of the above results, we propose to apply warm and cold sensations to interfaces in private cars or industrial fields, in order to give users or workers a more error-free, comfortable, and enjoyable experience. While the hardware technologies related to a touch sense are progressing, it is important to explore how such technologies should be utilized to satisfy users' psychological aspects. This study focused on matching users' intuitive images, and obtained the basic data that can be useful to realize users' intrinsic comfort and satisfaction.

References

[1] Stevens, J.C., Green, B.G.: Temperature-touch interaction, Weber's phenomenon revisited. Sensory Processes 2, 206–219 (1978)

[2] Green, B.G.: Localization of thermal sensation, An illusion and synthetic heat. Percept Psychophys 22, 331–337 (1977)

[3] Stevens, J.C.: Thermal intensification of touch sensation, Further extensions of the Weber phenomenon. Sensory Processes 3, 240–248 (1974)

[4] Stevens, J.C., Choo, K.K.: Temperature sensitivity of the body surface over the life span. Somatosens Mot. Res. 15, 13–28 (1998)

[5] Gibson, J.J.: Observation on active touch. Psychol. Rev. 69, 477–490 (1962)

[6] Gibson, J.J.: The senses considered as perceptual systems. Houghton, Mifflin (1966)

[7] Norman, D.A.: The design of everyday things. Doubleday, New York (1990)

[8] Katz, D.: Der Aufbau der Tastwelt. Barth, Leiptig (1925); Krueger, L.E. (ed.) The world of touch. Erlbaum, Hillsdale (1989)

[9] Nakanishi, M., Yamamoto, S.: What Images Are Evoked by Tactile Feeling? -For Designing Control Interfaces Taking Tactile Feeling into Consideration. In: Proceedings of the Workshop on Tactile and Haptic Interaction, Tokyo, Japan, pp. 22–29 (2007)

[10] Nakanishi, M., Okada, Y., Yamamoto, S.: A Study on Effective Tactile Feeling of Control Panels for Electrical Appliances. In: Proceedings of the 13th International Conference on HCI (Human Computer Interaction), San Diego, USA, on CD-ROM (2009)

Information Processing for Constructing Tactile Perception of Motion: A MEG Study

Ayumi Nasu[1], Kentaro Kotani[2], Takafumi Asao[2], and Seiji Nakagawa[3]

[1] Graduate School of Science and Engineering, Kansai University
3-3-35 Yamate-cho, Suita, Osaka 564-8680, Japan
[2] Faculty of Engineering Science, Kansai University
3-3-35 Yamate-cho, Suita, Osaka 564-8680, Japan
[3] Health Research Institute, National Institute of Advanced Industrial Science and Technology (AIST)
1-8-31 Midorigaoka, Ikeda, Osaka 563-8577, Japan
k236521@kansai-u.ac.jp

Abstract. We clarified whether temporal changes in information processing exist by comparing real and apparent motion conditions when tactile moving perception was given. We used magnetoencephalography (MEG), which has a high temporal resolution, to capture the dynamic changes in brain information processing. As a result, it was revealed that temporal transition through several cortexes was shown in which initial information processing occurred in the somatosensory cortex, followed by MT/V5, and then the activity was transmitted to motor-related areas when tactile moving stimuli were given.

Keywords: Magnetoencephalography, Tactile, Apparent motion, Real motion, Perception of speed, Somatosensory, Motor cortex, MT/V5.

1 Introduction

Recently, a study associated with tactile interface to present information at the skin surface has been conducted [1]. When characters are presented via tactile interface, a technique using moving stimuli is more effective than static presentation. By applying such a technique, apparent motion, a type of tactile illusion normally known as visual illusion, has been used [2]. Apparent motion is perceived as if a stimulus is continually moving from the initial location toward the subsequent location on the skin surface by presenting several stimuli at a certain temporal and spatial distribution [3]. Tactile apparent motion can be presented perception of different speed when stimulus duration and inter-stimulus onset interval (ISOI) are optimally changed [4]. Tactile apparent motion stimuli can generate perception of speed, which offers rich information representing emergency and priority of information via tactile moving stimuli [4]. For the development of tactile interface in the future, it is necessary to research the fundamental mechanisms of information processing associated with perception of tactile apparent motion in the brain [5] along with the psychological study to clarify tactile apparent motion.

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 478–487, 2011.

It has been under debates how is tactile information obtained on the skin surface processed in the brain. It is well known by previous studies that the activation occurs in the somatosensory area, corresponding to the location of stimulus presentation in the body when tactile stimuli were presented [6]. Moreover, it has been shown that the area MT/V5, which is related to visual motion perception, was activated when tactile real motion stimuli were presented to the skin surface [7]. A recent study of brain activity generated by tactile apparent motion stimuli applied to the fingertips showed activation not only in the somatosensory area but also in the motor area only when subjects perceived apparent motion stimuli compared with the simultaneous presentation of stimuli [8]. Although tactile apparent motion is given by a set of discrete single stimuli, perception of motion is obtained. Motion perception provided by tactile apparent motion implied the possibility of higher level information processing such that discretely given tactile stimuli were transferred and constructed into continuous moving perception. We expected that the area MT/V5 is activated not only when real moving stimuli were presented but also when the stimuli consisting of tactile apparent motion are presented. It has not been clarified how the perception of speed provided by tactile apparent motion stimuli was processed in the brain by now. We expected that this research is clarified any changes in information processing time in related areas of the brain when the tactile apparent motion conditions are given different velocity which included an element of time. We clarified whether or not temporal changes in information processing exist by comparing real and apparent motion conditions when tactile moving perception was given.

2 Methods

2.1 Subjects

A total of 13 healthy college students (age range: 20-23 years; right-handed) participated in the experiments. 11 participated in both apparent motion experiment and real motion experiment and 2 participated in only apparent motion experiment.

2.2 Tactile Stimuli Presentation Device

Tactile apparent motion was generated by air-jet. The equipment for presenting tactile apparent motions consisted of an air compressor, a precision regulator, an electro pneumatic regulator, and an electro-magnetic valve, each of which was connected serially to produce air-jet from the nozzle. Tactile real motion was presented at the skin surface by moving a wooden pin. The equipment for presenting tactile moving stimuli consisted of similar elements used for tactile apparent motions except for a wooden pin which was moved by compressed air generated via the nozzles (Fig. 1). The nozzle's internal diameter using the tactile apparent motion stimuli presentation device was 1.0 mm. The movable pin's diameter using the real moving tactile stimuli presentation device was 2 mm.

2.3 Stimulated Area

Subjects were seated with their right hand fixed on a table under which nozzles were located. A set of the stimulus condition of tactile apparent motion was given to the

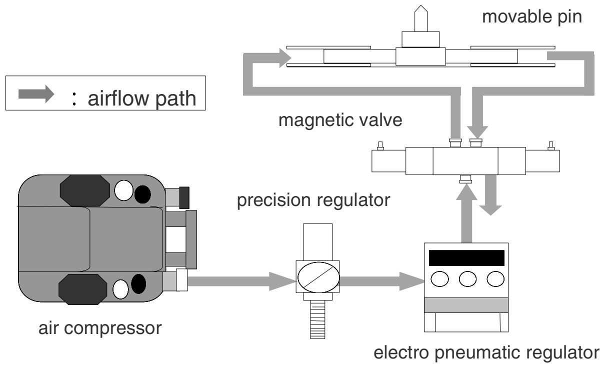


Fig. 1. Real moving tactile stimuli presentation device

subject's right index finger. The first stimulus was given between the distal interphalangeal (DIP) and metacarpal phalangeal (MP) joints, whereas the second stimulus was given between the proximal interphalangeal (PIP) and DIP joints. The distance between nozzles was 30 mm, and the interval between the nozzle and the skin was 5.0 mm (Fig. 2(a)). For the stimulus condition of real tactile motion, the stimulus started between the DIP and MP joints and moved 30 mm forward to the location at the middle point between the DIP and PIP joints (Fig. 2(b)).

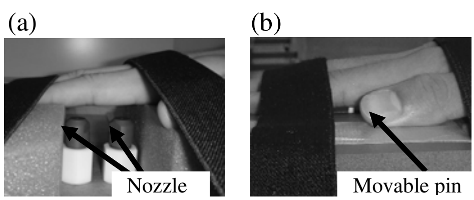


Fig. 2. (a) Tactile apparent motion stimulated area, (b) Real moving tactile stimulated area

2.4 Stimulation Paradigm

For the stimulus condition of tactile apparent motion, three levels of velocity perception were provided by changing duration and ISOI, that is, 50 ms of duration and 30 ms of ISOI for high-speed condition, 100ms of duration and 70ms of ISOI for middle-speed condition, and 200 ms of duration and 150 ms of ISOI for low-speed condition. A total of 60 stimuli for each stimulus condition were randomly presented. All trials were completed in approximately 11 minutes.

For the stimulus condition of real tactile motion, 60 trials were presented. All trials were completed in approximately 6 minutes.

2.5 Data Analysis

We used MEG due to its high temporal resolution for capturing dynamic changes in information processing in the brain. MEG data were recorded by a whole-head multi-channel DC-squid neuromagnetometer. Data were filtered on-line with a band-pass filter of 0.03-100 Hz and digitized at a sampling rate of 400 Hz. The signal was averaged off-line with epochs ranging from -500 to 1000 ms. Trigger signals for opening electromagnetic valves were used for the initial tactile stimulus as well as time onset for averaging the MEG signals for the stimulus condition of apparent motion. For the stimulus condition of real motion, the signal for moving the pin was obtained by an optical sensor. The signal was also used as a trigger signal. Averaged data were bandpass filtered between 1 and 30 Hz. Baseline was set to the averaged MEG signal using between -500 and -300 ms. From this averaged MEG data, root mean square (RMS) data were obtained to evaluate the intensity and latency of activation in related areas of the brain. Equivalent current dipoles (ECDs) were used to clarify the transition of information processing involving tactile motion stimuli.

3 Results

Figure 3 shows typical changes in the RMS values of MEG responses in three different regions of the brain for the high-speed condition of tactile apparent motion (subject G). Peak activities were observed at the left somatosensory area (LS) at 90.0 ms, left motor area (LM) at 382.5 ms, and left area MT/V5 (LMT) at 247.5 ms.

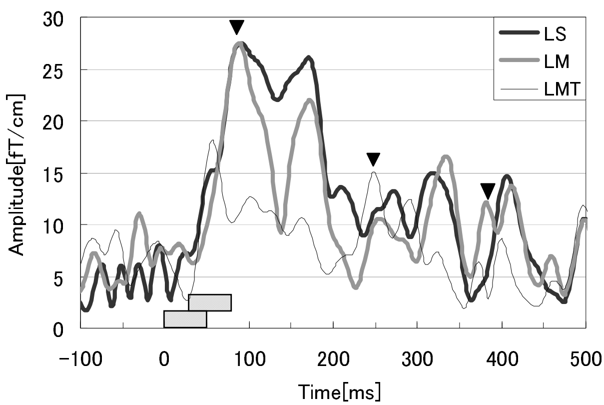


Fig. 3. RMS values of MEG responses in three different regions of the brain for the high-speed condition of tactile apparent motion (subject G). Gray blocks under the waveforms show the duration from stimulus onset.

Figure 4 shows typical changes in the RMS values of MEG responses in three different regions of the brain for the tactile real motion condition (subject G). Peak activities were observed at the LS at 32.5 ms, LM at 287.5 ms, and LMT at 237.5 ms.

Figure 5 shows the estimated sources of MEG responses when tactile motion perception was given (subject G, high-speed condition for tactile apparent motion). Each figure is superimposed on magnetic resonance images. When tactile motion perception was given, the sources of MEG responses were estimated at the LS at 85.0 - 92.5 ms (Fig. 7 (a)) and at the LM at 360.0 - 387.5 ms (Fig. 7 (b)).

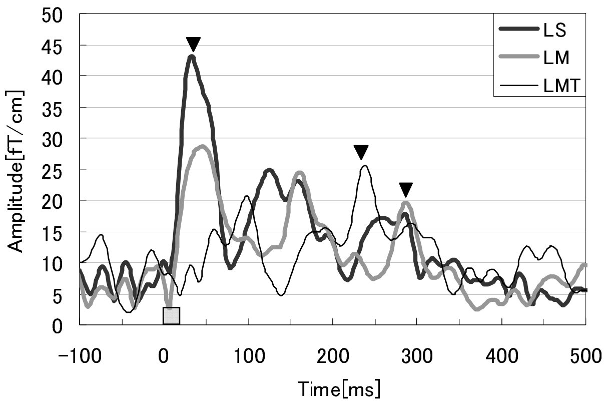


Fig. 4. RMS values of MEG responses in three different regions of the brain for the tactile real motion condition (subject G). Gray blocks under the waveforms show the duration from stimulus onset.

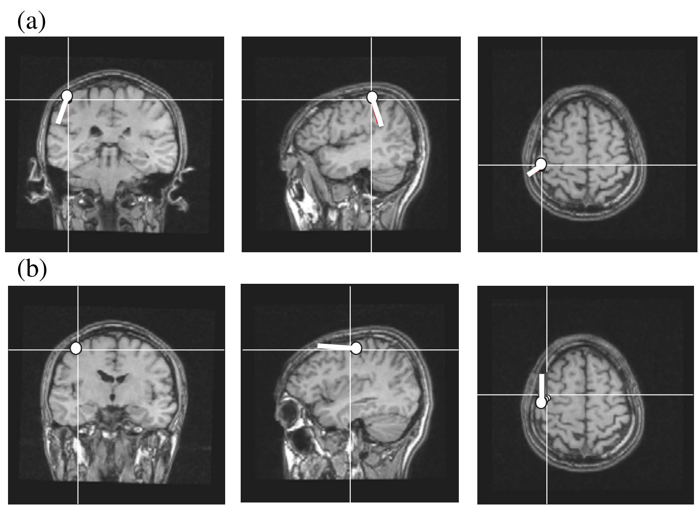


Fig. 5. Estimated sources of MEG responses with tactile motion perception (subject G, high-speed condition for tactile apparent motion). (a) LS (85.0 - 92.5 ms from the onset of the first stimulus presentation), (b) LM (360.0 – 387.5 ms from the onset of the first stimulus presentation). Acceptance criteria for ECD localization include (i) > 90% of goodness of fit, (ii) < 10000 mm^3 of confidence volume, (iii) and dipole moments significantly lasted at least 10 ms.

4 Discussion

4.1 Somatosensory Area

Previous studies observing brain activity when a single tactile stimulus was presented demonstrated activation of the contralateral primary somatosensory area at 20 - 60 ms following stimulus onset [8]. In the present results, the contralateral somatosensory area was activated when tactile apparent motion stimuli and real motion stimuli were given, which was in consistent with a previous study.

4.2 Motor Area

Activation in the motor area was previously demonstrated when tactile apparent motion stimuli was given, suggesting that the motor area might function for high-order tactile motion perception [8]. Furthermore, the premotor area was reported to be strongly associated with the dorsolateral prefrontal cortex, which is responsible for "attention" and "memory" [9]. Hence, it was supposed that subjects attended to the stimuli, evoked sense of stroking stimuli from their essential memory and perceived tactile motion.

4.3 Area MT/V5

The area MT/V5 is reported to process not only information regarding visual real motion but also apparent motion [10]. In the present study, the area MT/V5 was activated not only by tactile real motion stimuli but also by tactile apparent motion stimuli. Presently, results showing recruitment of visual areas in a supposedly tactile task are open to an alternative interpretation in terms of visual imagery (e.g., activation in the area MT/V5 may be evoked by visually imagining movement, rather than being specifically associated with the processing of tactile motion) [7]. We assumed that the area MT/V5 was activated because many subjects visually interpreted a set of tactile stimuli as a movement. Additionally, a previous study of brain activity with tactile single stimuli demonstrated no activation in the area MT/V5 [7]. We consider that the activation of the area MT/V5 with tactile stimuli is one supporting evidence that subjects perceived and constructed the motion of stimuli.

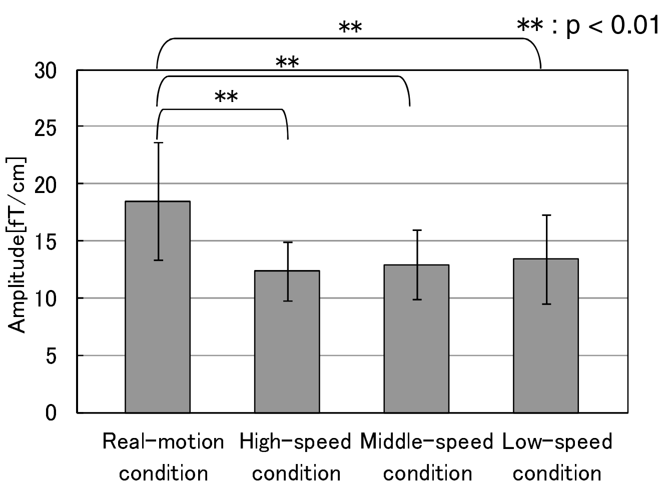


Fig. 6. Mean peak amplitude with standard deviations in the left area MT/V5 in each condition. The different significance levels of the paired, one-tailed t test are shown with **: $p < 0.01$.

4.4 Comparisons of the Peak Amplitude between Conditions of Tactile Apparent Motion and Tactile Real Motion

As shown in Fig. 6, the mean peak amplitude in the real motion condition was significantly ($p < 0.01$) higher than that in all other conditions for tactile apparent motion in the left area MT/V5.

Comparing the tactile real motion condition with all the conditions of tactile apparent motion, the peak amplitude in the tactile real motion condition was significantly higher than that in all the conditions of tactile apparent motion in the left area MT/V5. A previous study of brain activity involving visual perception using visual apparent motion stimuli and visual real motion stimuli in which various distances of stimulus locations were changed, demonstrated that the peak amplitude of real motion stimuli was higher than that of apparent motion stimuli [11]. Moreover, significant differences in the peak amplitudes in the left area MT/V5 between conditions were observed because subjects could visualize movement easier than that of the tactile apparent motion stimuli.

4.5 Changes in the Mean Time Intervals from the Peak in the Left Somatosensory Area to the Peak in the Left Area MT/V5 with Changes in the Perceived Velocity Generated by a Set of Stimuli

As shown in Fig. 7, the mean time intervals from the peak latency in the left somatosensory area to the peak latency in the left area MT/V5 between all conditions showed a significant difference ($p < 0.01$). In the tactile apparent motion condition, as perceived speed generated by the stimuli became higher, the mean time intervals from the peak latency in the left somatosensory area to that in the left area MT/V5 became shorter. The duration generated by the condition for tactile real motion stimuli was approximately 20 ms higher than that generated by the high-speed condition for tactile apparent motion. This demonstrated that the mean time interval obtained by the real motion condition was much shorter than that obtained by the high-speed condition.

Kawakami et al. [12] showed that the peak latency of the area MT/V5 became shorter when perceived velocity generated by the visual stimuli became higher. In the present study, as perceived velocity generated by the tactile stimuli became higher, the mean time intervals from the peak time in the left somatosensory area to that in the left area MT/V5 became shorter (Fig. 7). It was considered that the duration to transmit information from the left somatosensory area to the left area MT/V5 became shorter as perceived velocity generated by the tactile stimuli became higher. In our experimental conditions, the duration of stimuli application became longer as perceived speed generated by the stimuli became lower. Thus, a time lag occurred from the receipt of sensory information generated by tactile stimuli in the somatosensory area to the start of information processing in the other related areas by changing the duration of stimulus presentation. Unfortunately in our experimental conditions, the perceived velocity generated by the stimuli was proportional to the time intervals from the peak in the somatosensory area to the peak in the area MT/V5. This was because the tactile apparent motion was an illusion phenomenon so that the perceived velocity given by tactile apparent motion could not be evaluated quantitatively. However, further implementation of devices associated with variable velocity changes in real pin movement would clarify such proportional relationships. There was also a possibility of a different relationship between the perceived velocity and the time intervals from the peak in the somatosensory area to the peak in the area MT/V5 when the tactile apparent motion stimuli are compared with the real moving tactile stimuli. It would be a future challenge to control the condition of the experiment using tactile apparent motion stimuli and real moving tactile stimuli such that both stimuli generate identical magnitudes of perceived velocity.

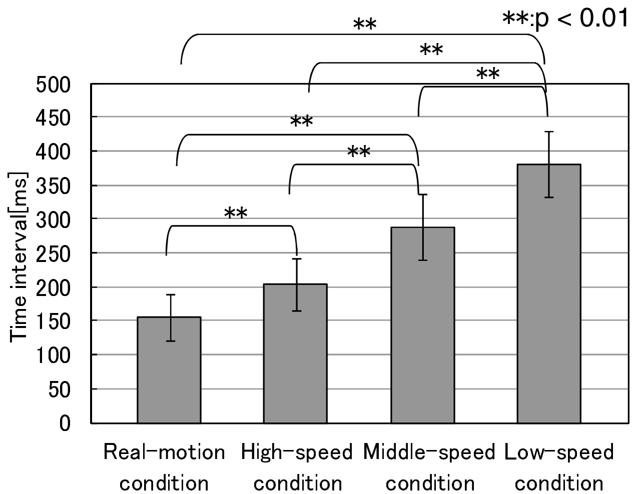


Fig. 7. Mean time intervals from the peak latency in the left somatosensory area to the peak latency in the left area MT/V5 in each condition and standard deviations. Paired, one-tailed t test; **: $p < 0.01$.

4.6 Information Processing for Tactile Motion Perception

Figure 8 shows the time chart of brain information processing for the tactile apparent motion stimuli of three velocity conditions as well as the condition of real moving tactile stimuli.

As shown in Fig. 8, the time intervals between the duration from stimulus onset and the peak latency of the motor area were relatively constant, even though the type of motion was different. In the tactile apparent motion stimuli, the subjects started to perceive stimulus movement after the second stimulus was given. In the real moving tactile stimuli, the subjects perceived stimulus movement once the stimulus was started. Therefore, there may be a constant time interval from realizing the existence of stimulus movement to processing, in a supplemental manner, a high-order tactile motion perception in the motor area.

Inter-cortical information processing for constructing tactile motion perception can be summarized as follows. The somatosensory area initially processes sensory information given by tactile stimuli. Then, when continuous or “semi-continuous” sets of stimuli, as represented by apparent motion, are received, the high-order visual area MT/V5 processes the visual images of tactual perception of motion. Furthermore, the time interval between the peaks of the somatosensory cortex and at the motor cortex suggested that the motor area processed information to supplement a high-order tactile motion perception and to evoke sense of stroking stimuli assisted by memory.

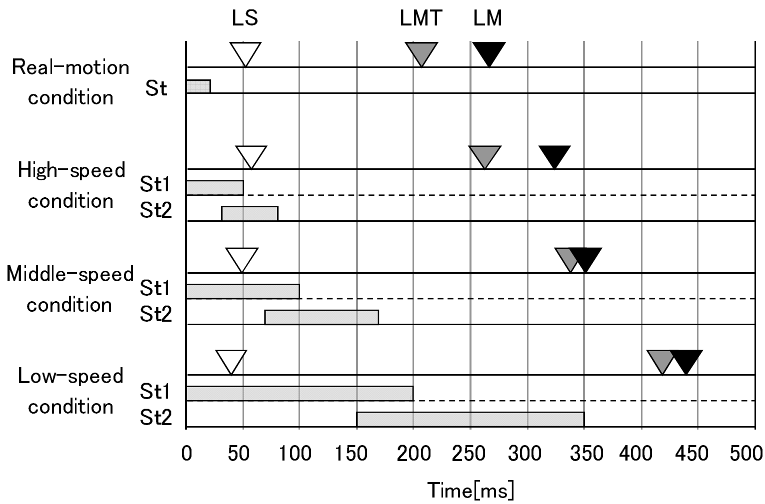


Fig. 8. Information processing for tactile motion perception. White inverted triangles indicate the activation time of the LS. Gray inverted triangles denote the activation time of the LM. Black inverted triangles represents the activation time of the LMT. Gray blocks under the time chart show the duration from stimulus onset in each condition. St1 denotes the first stimulus of the tactile apparent motion, St2 indicates the second stimulus, and St represents the tactile real motion stimulus.

5 Conclusion

The results associated with dipole moment locations revealed that the contralateral motor area was activated when tactile motion perception was given, consistent with a previous study [8]. The high-order visual area MT/V5 was activated by not only tactile real motion stimuli but also tactile apparent motion stimuli. When tactile motion perception was given, sensory information generated by the tactile stimuli was transmitted to the area MT/V5 and then to the motor area after completing the initial processing in the somatosensory area. As perceived velocity generated by the tactile stimuli became lower, a time lag was apparent from the receipt of sensory information generated by tactile stimuli in the somatosensory area to the start of information processing in the other related areas.

Acknowledgements. This work was partially supported by JSPS Kakenhi (20370101).

References

1. Hafez, M.: Tactile interfaces: technologies, applications and challenges. The Visual Computer 23, 267–272 (2007)
2. Shimizu, Y., Saida, S., Wake, T., Nakamura, A., Ohzu, H.: Optimum design of tactile display for a reading aid: Uses of computers in aiding the disabled, pp. 383–391. North-Holland Publishing, Amsterdam (1982)
3. Kirman, J.H.: Tactile apparent movement: The effects of interstimulus onset interval and stimulus duration. Perception & Psychophysics 15, 1–6 (1974)
4. Kohli, L., Niwa, M., Noma, H., et al.: Towards effective information display using vibrotactile apparent motion. In: Proceedings of the 2006 IEEE Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, pp. 445–452 (2006)
5. Hayward, V.: A brief taxonomy of tactile illusions and demonstrations that can be done in a hardware store. Brain Research Bulletin 75, 742–752 (2008)
6. Hari, R., Reinikainen, K., Kaukoranta, E., et al.: Somatosensory evoked cerebral magnetic fields from SI and SII in man. Electroencephalogr. Clin. Neurophysiol. 57, 254–263 (1984)
7. Hagen, M.C., Franzen, O., McGlone, F., et al.: Tactile motion activates the human middle temporal/V5(MT/V5) complex. Eur. J. Neuroscience 16, 957–964 (2002)
8. Karageorgiou, E., Koutlas, I., Alonso, A., et al.: Cortical processing of tactile stimuli applied in quick succession across the fingertips: temporal evolution of dipole sources revealed by magnetoencephalography. Exper. Brain. Res. 189, 311–321 (2008)
9. Simon, S.R., Meunier, M., Piettre, L., et al.: Spatial attention and memory versus motor preparation: Premotor cortex involvement as revealed by fMRI. J. Neurophys. 88, 2047–2057 (2002)
10. Kaneoke, Y., Bundou, M., Koyama, S., et al.: Human cortical area responding to stimuli in apparent motion. NeuroReport 8, 677–682 (1997)
11. Yang, C.Y., Hsieh, J.C., Chang, Y.: An MEG study into the visual perception of apparent motion in depth. Neuroscience Lett. 403, 40–45 (2006)
12. Kawakami, O., Kaneoke, Y., Maruyama, K., et al.: Visual detection of motion speed in humans: Spatiotemporal analysis by fMRI and MEG. Human Brain Mapping 16, 104–118 (2002)

A Study on Selection Ability in the 3D Space by the Finger

Makoto Oka, Yutaro Ooba, Hidetaka Kuriiwa, Ryuta Yamada, and Hirohiko Mori

Tokyo City University, 1-28-1 Tamadutumi, Setagaya, Tokyo, Japan
{moka,hmori}@tcu.ac.jp, {ooba,kuriiwa,yamada}@ims.tcu.ac.jp

Abstract. Intuitive gestures are very effective for interactions. Pointing gesture with a finger would be used for interactions like pie menu selection. It has been researched as to the appropriate numbers of menu items being displayed in a pie menu. However, no research has been made for the case where menus are selected by using gestures. We will experiment in order to examine the ability of pointing gesture (selection ability) in the 3-dimensional space. The experiment was conducted in order to find out the test subjects' selection ability of pointing gesture. By combining the obtained resolution abilities of pointing gesture at the 2-dimentional surface of Pitch and Yaw, it is suggested that the selection ability of selection behavior in the 3-dimentional space is 22 areas.

Keywords: Gesture, Selection Ability, Menu, Pie menu.

1 Introduction

Intuitive gestures are very effective for interactions. In recent years, the researches on interactions, which operate various products and services such as music players and TV by using gestures, are being actively conducted.

It is difficult for computers to recognize the intentions of users from their natural gestures. Although it is relatively easy for computers to recognize standard gestures like sign language (commands), users need educations and trainings to be able to use them. Both have merits and demerits. Generally speaking, gestures are very effective for interactions if the use is limited to interactions like making selections from the displayed menu[1].

However, as the case now stands, there is no coherence between gestures and the allocations of interactions, and each application and service makes its own allocations. Products and services have become diversified and multifunctional, and there are more functions and interactions which need to be allocated to gestures, thereby the allocation has become difficult.

Therefore, pointing gesture with a finger would be used for interactions like menu selection. It is possible to select with simple gesture even when there are many menu items. In addition, pointing gesture is intuitive, thus users need no special education or training.

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 488–497, 2011.

Pie menu is a kind of GUI menus. It displays menu items in a radial fashion centering on the mouse pointer. Users select a menu item by the direction to which they move the mouse pointer. Pie menu can be selected by the angles. It is suitable for interactions using pointing gesture.

It has been researched as to the appropriate numbers of menu items being displayed in a pie menu. However, no research has been made for the case where menus are selected by using gestures. In order to find it out, it is necessary to know human ability of pointing properly. Therefore, we will experiment in order to examine the ability of pointing gesture (selection ability) in the 3-dimensional space.

2 Related Research

Miyamoto et al. have discussed how they should allocate menu items to the extended pie menu[2]. They examined how the selection accuracy varied depending on the sizes and locations of each area within the pie menu so that they could discuss the most appropriate area pattern of the pie menu.

Saeki et al. have attempted to examine the selection ability of pointing gesture in the 3-dimentional space[3]. The experiment was conducted without posing restrictions on the arm joints of the test subjects, which allowed the free movements. Because each test subject had different pointing gestures, they were not able to conduct statistical analysis effectively.

Therefore, we thought it was necessary to examine the range of motion and the arthresthesia of joints of shoulder, elbows, wrists and fingers separately.

3 Experiment 1

The experiment was conducted in order to find out the test subjects' selection ability of pointing gesture. With their forearms fixed horizontally, the selection ability for Pitch (vertical) direction was obtained. The selection ability for Yaw (horizontal) direction was obtained for the each obtained selection ability for Pitch (Fig.1.).

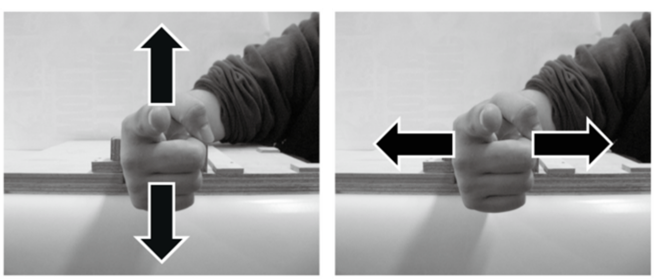


Fig. 1. Pointing Gesture by the finger(Left: Pitch, Right: Yaw)

3.1 The Experiment to Obtain the Selection Ability for Pitch

The below restrictions were added to the vision and arm joints of the test subjects

1. Visual feedback was eliminated.
2. The joints of arms and fingers, which were except for wrists, were fixed. (Fig.2.).
3. The joints of wrists were fixed in a way to prevent adduction and abduction.

Pointing gesture in this research is by using the wrist and finger joints. In order to separate pointing gestures into Pitch (vertical direction) and Yaw (horizontal direction), the condition (2) was added. With the forearm fixed, the hemisphere centering on the direction from the elbow to the wrist was considered. The test subjects were instructed to point at the center of a certain area, which was a part of the partitioned hemisphere. The accuracy of their pointing gestures was examined.

In order to obtain the pointing angles of the test subjects, 3D sensor module (TDS01V[4]) was attached on top of their index finger. TDS01V is a sensor equipped with a triaxial acceleration speed sensor and a triaxial geomagnetic sensor.

The test subjects were 18 males and females who were in early twenties and right-handed. All of them had normal wrists joints. It was defined that one trial consisted of the test subjects pointing at the center of an area as instructed one time. It was defined that one set consisted (135 trials) of one trial for each 2 to 16 partitioned areas of Pitch. Two sets were completed in the experiment. The order of the trials was random within one set. Each set had different order. The experiment took approximately 60 minutes.

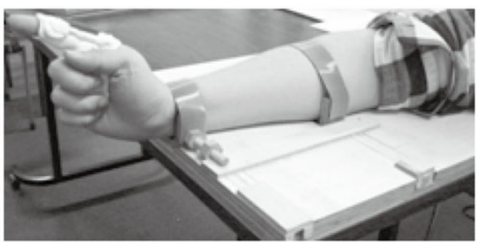


Fig. 2. The joints of arms and fingers were fixed

Evaluation Method. The results of measuring the pointing angles of the gestures were analyzed. The pointing angles in one particular area of all the test subjects were nearly normal distribution. We let the average pointing angle in a particular area be μ and its standard variation be σ, assuming that they were able to point with the accuracy of approximately 68.3% if $\mu\pm\sigma$ was within the area. they were able to point with the accuracy of approximately 95.5% if $\mu\pm2\sigma$ was within the area. they were able to point with the accuracy of approximately 99.7% if $\mu\pm3\sigma$ was within the area.

Because the wrists have a range of motion, it is not possible to point at outside the range. In the case of Pitch, the area above the center is outside the range of motion. It was necessary to evaluate by excluding such area.

Results. The results of pointing for Pitch are shown in Table 1. From Table 1, it is possible to obtain the selection ability of pointing gesture at the 2-dimentional surface. The column of $\mu \pm \sigma$ of Table 1 shows whether $\mu \pm \sigma$ from the distribution of the pointing angles were within a particular area. If so, it indicates that it was possible to properly point with that accuracy. It is same for the column of $\mu \pm 2\sigma$ and that of $\mu \pm 3\sigma$. check cell means possible to point, while blank cell means impossible to point properly.

Table 1. The accuracy of pointing gesture for Pitch

Number of partitions	Area No.	Central angle (degrees)	$\mu \pm \sigma$	$\mu \pm 2\sigma$	$\mu \pm 3\sigma$
2	1 (Up)	-45.0			
	2 (Bottom)	45.0	✓	✓	✓
3	1	-60.0			
	2	0.0	✓	✓	✓
	3	60.0	✓	✓	
4	1	-67.5			
	2	-22.5			
	3	22.5	✓		
	4	67.5	✓		
5	1	-72.0			
	2	-36.0			
	3	0.0	✓		
	4	36.0	✓		
	5	72.0	✓		

As to $\mu \pm 3\sigma$, it was possible to properly point at the area 2, which were parts of the three-partitioned area. As to $\mu \pm 2\sigma$, it was possible to properly point at the area 2 and 3 which were parts of the three-partitioned area. As to $\mu \pm \sigma$, it was possible to properly point at the area 3, 4 and 5, which were parts of the five-partitioned area.

Discussion. In one trial of pointing gesture, deviation was developed between the actual angle of pointing gesture and the angle which was meant to be. The scatter diagram of the deviations shows the tendency that the actual angles deviated from the correct angles toward the center (toward the front). It was indicated that there was a certain tendency of deviation between the correct angles and the actual pointing angles. By doing a regression analysis, corrected values were obtained based on the regression expression. The corrected angle y of Pitch is Eq. 1.

$$y = -0.195x + 9.948 \tag{1}$$

The accuracy of pointing after the correction is shown in Table 2. As a result of the correction, the selection ability for Pitch was not improved at to $\mu \pm \sigma$. It was found out possible to properly point at the area 3 (the central angle 0 degree), the area 4 (-36 degrees) and the area 5 (-72 degrees) out of the five-partitioned area of Pitch.

Table 2. The accuracy of pointing gesture for Pitch after the correction

Number of partitions	Area No.	Central angle (degrees)	μ±σ	μ±2σ	μ±3σ
2	1	-45.0			
	2	45.0	✓	✓	✓
3	1	-60.0			
	2	0.0	✓	✓	✓
	3	60.0	✓		
4	1	-67.5			
	2	-22.5			
	3	22.5	✓		
	4	67.5	✓		
5	1	-72.0			
	2	-36.0			
	3	0.0	✓		
	4	36.0	✓		
	5	72.0	✓		

Table 3. The accuracy of pointing gesture for Yaw (0 degrees) after the correction

Number of partitions	Area No.	Central angle (degrees)	μ±σ	μ±2σ	μ±3σ
2	1 (Left)	-45.0	✓	✓	✓
	2 (Right)	45.0	✓	✓	
3	1	-60.0	✓		
	2	0.0	✓	✓	
	3	60.0	✓		
4	1	-67.5	✓		
	2	-22.5	✓		
	3	22.5	✓		
	4	67.5	✓		
5	1	-72.0	✓		
	2	-36.0	✓		
	3	0.0	✓		
	4	36.0			
	5	72.0	✓		

3.2 The Experiment to Obtain the Selection Ability for Yaw

It was experimented in order to obtain the selection ability for Yaw at the area of 3 and 4 out of the five-partitioned area of Pitch. As to the area 5, it was not experimented since the range of motion of wrists was extremely small.

The experiment was conducted under each condition where the movable surface of the wrists was fixed at 0 (horizontal) and -36 degrees. The experiment conditions and evaluation method were same as Pitch in the experiment 1.

Results. There was a tendency that the pointing angles uniformly deviated toward the same direction. Table 3 and Table 4 show that the accuracy of pointing which was obtained by using corrected data.

From Table 3, as to μ±3σ, it was possible to properly point at the area 1, which were parts of the two-partitioned area. As to μ±2σ, it was possible to properly point at the area 1 and 2, which were parts of the two-partitioned area. As to μ±σ, it was possible to properly point at the area 1, 2, 3 and 4, which were parts of the four-partitioned area.

Table 4. The accuracy of pointing gesture for Yaw (-36 degrees) after the correction

Number of partitions	Area No.	Central angle (degrees)	μ±σ	μ±2σ	μ±3σ
2	1	-45.0	✓	✓	✓
	2	45.0	✓	✓	✓
3	1	-60.0	✓		
	2	0.0	✓	✓	
	3	60.0	✓	✓	
4	1	-67.5	✓		
	2	-22.5	✓		
	3	22.5	✓		
	4	67.5			
5	1	-72.0	✓		
	2	-36.0	✓		
	3	0.0	✓		
	4	36.0	✓		
	5	72.0			

From Table 4, as to μ±3σ, it was possible to properly point at the area 1 and 2, which were parts of the two-partitioned area. As to μ±2σ, it was possible to properly point at the area 1 and 2, which were parts of the three-partitioned area. As to μ±σ, it was possible to properly point at the area 1, 2, 3 and 4, which were parts of the five-partitioned area.

3.3 Discussion

In one trial of pointing gesture, deviation was developed between the actual angle of pointing gesture and the angle which was meant to be. The analysis was made based on the idea that non-overlapping of adjacent areas meant that they were able to point properly, rather than obtaining the selection ability based on whether they were able to point at where they intended.

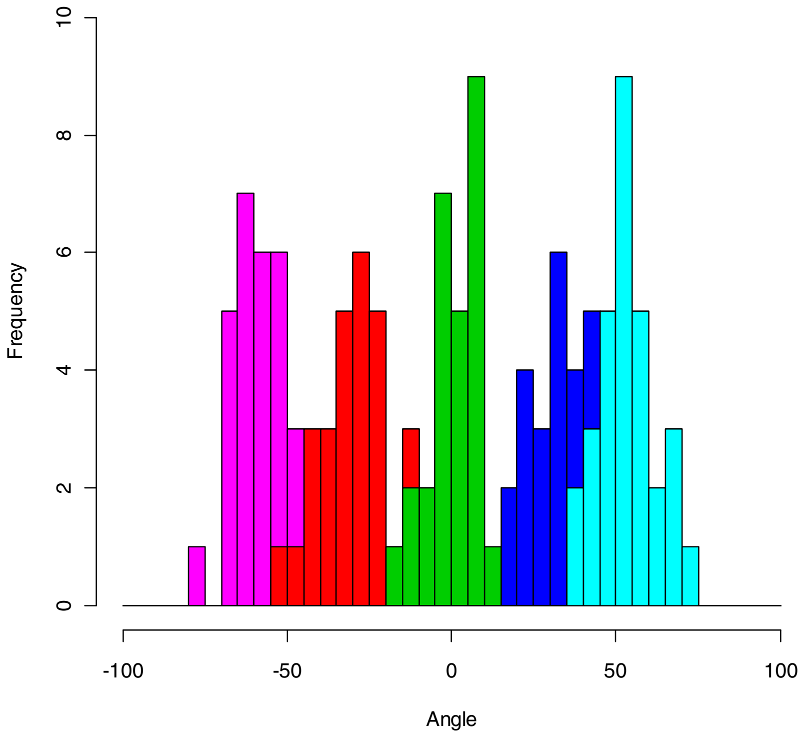


Fig. 3. Histogram of pointing gesture for Pitch of five-partitioned area

Figure 3 is the histogram of the results of partitioning the area into five toward the Yaw direction and at 0 degree of Pitch direction (horizontal surface). It is suggested that they can properly point more areas than four partitions. It is necessary to experiment while considering the analysis methods.

4 Experiment 2

It was experimented by removing the condition 3 of the experiment 1 and adding the adduction and abduction of wrists to the experiment 1. The experiment conditions and evaluation method were same as experiment 1. The experimental order was same as the experiment 1. The test subjects were 20 males and females who were in early twenties and right-handed. All of them had normal wrists joints.

4.1 The Experiment to Obtain the Selection Ability for Pitch

As the experiment 1, there was a tendency that the pointing angles uniformly deviated toward the same direction. By doing a regression analysis, corrected values were obtained based on the regression expression. Table 5 show the accuracy obtained by using the corrected data.

Table 5. The accuracy of pointing gesture for Pitch after the correction

Number of partitions	Area No.	Central angle (degrees)	μ±σ	μ±2σ	μ±3σ
2	1	-45.0	✓	✓	✓
	2	45.0	✓	✓	
3	1	-60.0	✓	✓	
	2	0.0	✓	✓	✓
	3	60.0	✓		
4	1	-67.5	✓		
	2	-22.5	✓	✓	
	3	22.5	✓		
	4	67.5	✓		
5	1	-72.0	✓		
	2	-36.0	✓		
	3	0.0	✓	✓	
	4	36.0	✓		
	5	72.0	✓		

Table 6. The accuracy of pointing gesture for Yaw after the correction

Number of partitions	Area No.	Central angle (degrees)	-72deg μ±σ	-36deg μ±σ	0deg μ±σ	36deg μ±σ	72deg μ±σ
2	1	-45.0	✓	✓	✓	✓	✓
	2	45.0	✓	✓	✓	✓	
3	1	-60.0	✓		✓		✓
	2	0.0	✓	✓	✓	✓	✓
	3	60.0	✓		✓	✓	
4	1	-67.5	✓		✓		
	2	-22.5		✓	✓	✓	
	3	22.5	✓	✓	✓	✓	✓
	4	67.5			✓		
5	1	-72.0					
	2	-36.0		✓	✓		
	3	0.0	✓	✓	✓		
	4	36.0			✓		
	5	72.0			✓		

As to $\mu\pm\sigma$, it was possible to properly point at the area 1, 2, 3, 4 and 5, which were parts of the five-partitioned area.

4.2 The Experiment to Obtain the Selection Ability for Yaw

It was possible to point properly up to five-partitioning of Pitch. It was experimented to examine the selection ability for Yaw as to the entire areas of five partitions of Pitch. The experiment was conducted under each condition where the movable surface of the wrists was fixed at -72, -36, 0 (horizontal), 36 and 72 degrees.

Results. The corrected measured values are shown in Table 6.

Discussion. It is the areas on the extreme left and the extreme right out of the entire area that showed a significant deviation of the observed values. It was found out that they became able to point properly at some areas after the correction. However, as to the extreme left of Yaw in the area 2 and area 3 in the five partitions of Pitch, they remained unable to point properly. It may be necessary to add other restrictions to pointing gesture.

5 Conclusion

In the experiment 1, they pointed with wrists joints under the condition where the arm joints were fixed and adduction and abduction of the wrists were restricted. By combining the obtained resolution abilities of pointing gesture at the 2-dimentional surface of Pitch and Yaw, it is suggested that the selection ability of selection behavior in the 3-dimentional space is 9 areas. The details of the 9 areas are the area 3, 4 and 5 out of five partitions of Pitch. The selection ability for Yaw in the area 3 of the five partitions of Pitch was four areas. The selection ability for Yaw in the area 4 of the five partitions of Pitch was four areas. The selection ability for Yaw in the area 5 of the five partitions of Pitch was one area. Combining those results indicates that the selection ability of selection behavior in the 3-dimentional space is 9 areas.

The experiment 2 was conducted by adding adduction and abduction of the wrists to the experiment 1. By combining the obtained resolution abilities of pointing gesture at the 2-dimentional surface of Pitch and Yaw, it is suggested that the selection ability of selection behavior in the 3-dimentional space is 22 areas.

6 Future Work

Since the selection ability obtained from the results had the added restrictions, they are not the true selection ability of free pointing gesture. As it is obvious from the results of this research report, the selection ability is improved as the degree of freedom of joints increases. It is necessary to experiment by changing the experimental conditions in the future to obtain selection ability of pointing gesture with more freedom.

In addition, the experiment methods used in this research were meant to obtain the 3-dimentional selection ability by combining those at 2-dimentional surface. It is necessary to further examine by using pointing gestures in the 3-dimentional space.

References

1. Lenman, S., Bretzner, L., Thuresson, B.: Using Marking Menus to Dvelop Command Sets for Computer Vision Based hand Gesture Interface. In: NordiCHI, pp. 239–242 (2002)
2. Miyamoto, M., Terada, T., Tsukamoto, M.: An Area Allocation Algorithm for Hierarchical Pie Menu. In: Human Interface 2009, pp. 15–22 (2009)
3. Saeki, H.: A Study on Selection Ability in the 3D Space by the Finger, Graduation thesis of Musashi Institute of Technology (2008)
4. TDS01V's pdf datasheet, `http://www.vitec.co.jp/`

Characteristics of Comfortable Sheet Switches on Control Panels of Electrical Appliances: Comparison Using Older and Younger Users

Yasuhiro Tanaka[1], Yuka Yamazaki[2], Masahiko Sakata[3], and Miwa Nakanishi[1]

[1] Keio University, Fac. of Science & Technology, Dept. of Administration Engineering
Hiyoshi 3-14-1, Kohoku, Yokohama 223-8522, Japan
tanakayasu63-425@a8.keio.jp, 1miwa_nakanishi@ae.keio.ac.jp
[2] Mitsubishi Electric Corp.,
5-1-1, Ofuna, Kamakura, 247-8501, Japan
Yamazaki.Yuka@cs.MitsubishiElectric.co.jp
[3] Mitsubishi Electric Group, Central Melco Corp.,
3-26-33, Takanawa, Minato, Tokyo, 108-0074, Japan
kiai106@ct-melco.co.jp

Abstract. This study focuses on sheet switches, which are among the most common user interfaces in electrical appliances, and explores the comfort level required by elderly users. Touching a switch is a common action in our daily lives; however, little research has been performed on this action. In particular, we were unable to find any reports on the comfort elderly users experience upon touching a sheet switch. Thus, electrical-appliance designers have no well-prepared reference data when designing new products. As a consequence, elderly users experience discomfort when using new appliances. Our goal is to construct a guideline for designing comfortable sheet switches for elderly users. In this study, as the first step toward achieving this goal, we attempt to clarify the physical parameters of sheet-switch features that contribute toward ease of use and comfort.

Keywords: Sheet switches, Electrical appliances, Elderly users, EMG, Comfort.

1 Introduction

One of the most famous research institutes in Japan, the Mitsubishi Research Institute (MRI), predicts that in 2030, the Japanese market of electrical appliances for the elderly will have expanded to one and a half times of that in 2010 [1]. Such electrical appliances have already been labeled as "mature products," and their focus is shifting from functions and performance to user comfort. Thus, designing user interfaces for such appliances by considering the physical, cognitive, and psychological characteristics of the elderly is becoming increasingly important for market growth in the near future.

At present, the user interfaces of commonly used electrical appliances such as rice cookers, washing machines, and microwave ovens are generally control panels

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 498–507, 2011.

consisting of few switches. In particular, switches covered with plastic sheets, “sheet switches,” are included in most recently designed products. Sheet switches have the advantage of flexibility in tactile and haptic design; therefore, they are expected to be applied to more products in the future. Our study focuses on such sheet switches, which are among the most common user interfaces of electrical appliances, and explores the comfort required by elderly users.

Little research has been performed on actions as common as touching a switch; in particular, we were unable to find any specific report on user comfort for the elderly when touching sheet switches. Thus, electrical-appliance designers face a challenge when designing new products because of the lack of well-prepared reference data. As a consequence, elderly users experience discomfort when using new appliances.

The Research Institute of Human Engineering for Quality Life (HQL) in Japan revealed that older users tend to push hard keys more strongly than younger users [2]. Although hard keys are no longer used in recently designed household products, this data suggests that a difference in practice exists between older and younger users when performing such simple actions, leading to the anticipation that a “comfortable” sheet switch may vary in application between these two groups. Our goal is to construct a guideline for designing comfortable sheet switches for the elderly. In this study, as the first step toward this goal, we attempt to clarify the physical parameters of sheet-switch features that contribute toward ease of use and comfort.

2 Experiment

2.1 Participants

Six healthy men and women 73 to 81 years old, and five healthy men and women 21 to 23 years old participated in the experiment, representing older and younger users, respectively. Both groups operated commonly used electrical appliances, such as rice cookers, washing machines, and microwave ovens, each day of the study.

2.2 Experimental Parameters

The following three parameters were expected to be related to comfort when touching the sheet switch:

(1) Size: defined as the diameter of the sheet switch (unit: mm).
(2) Depth: defined as the length from the top of the switch to the reaction point (unit: mm).
(3) Reaction force: defined as the force needed to complete a switch touch without plastic sheet (unit: N).

Three different levels were assigned to each of the above parameters and 27 types of sheet switches were designed as model switches for the experiment. The levels assigned to each parameter are summarized in Table 1. Each model switch was installed on an aluminum base as shown in Fig. 1.

Table 1. Levels assigned to each parameter of model switches

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
Size[mm]	28.0	28.0	28.0	28.0	28.0	28.0	28.0	28.0	28.0
Depth[mm]	0.4	0.9	1.5	0.5	0.8	1.4	0.5	1.0	1.5
Reaction force[N]	1.1	1.0	1.1	1.7	1.7	1.6	2.5	2.5	2.6
	[10]	[11]	[12]	[13]	[14]	[15]	[16]	[17]	[18]
Size[mm]	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0
Depth[mm]	0.4	0.9	1.4	0.5	1.0	1.5	0.5	1.0	1.5
Reaction force[N]	1.2	1.1	1.0	1.7	1.7	1.7	2.6	2.5	2.6
	[19]	[20]	[21]	[22]	[23]	[24]	[25]	[26]	[27]
Size[mm]	8.0	8.0	8.0	8.0	8.0	8.0	8.0	8.0	8.0
Depth[mm]	0.5	1.0	1.5	0.5	1.0	1.5	0.5	0.9	1.5
Reaction force[N]	1.1	1.0	1.1	1.7	1.7	1.7	2.5	2.6	2.5

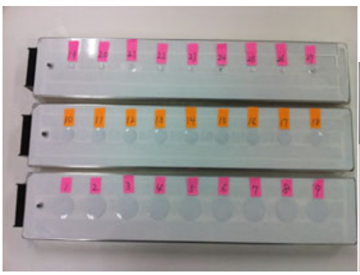


Fig. 1. Model switches

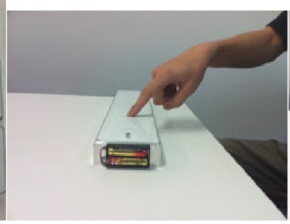


Fig. 2. Vertical touch

Fig. 3. Horizontal touch

2.3 Experimental Task

The participants operated each of the 27 model switches in the vertical and horizontal manner as shown in Figs. 2 and 3, respectively.

To examine vertical touch, users were instructed to approach the model switch in the same manner as they would when using a rice cooker in their usual daily lives. Similarly, a microwave oven was employed to examine horizontal touch. The model switches representing vertical touch were tested first and those representing horizontal touch were tested next. The order of testing trials was randomized to eliminate order effects.

2.4 Experimental Environment

In normal practice, electrical appliances such as a rice cockers and microwave ovens are operated from a standing position. However, considering that the participants included the elderly who required more than one hour to complete the testing trials, we prepared the experimental environment such that the tasks could be performed from a sitting position. According to some electrical-appliance manufacturers, rice cockers and microwave ovens are designed considering the assumption that the appliance is placed at a height of approximately 100 cm from the ground, which is indeed a common practice in many homes. Thus, considering that the height of users' arms change by approximately 30 cm between standing and sitting positions, the model switches were placed 70 cm above the ground.

2.5 Data

The participants evaluated the comfort they experienced upon touching each model switch by assigning scores ranging from −2 to +2. An answer such as "between +1 and +2" was allowed and recorded as a score of +1.5.

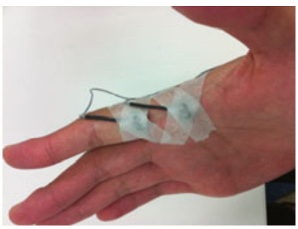


Fig. 4. Lumbrical muscles

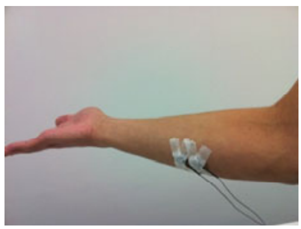


Fig. 5. Flexor carpi ulnaris muscles

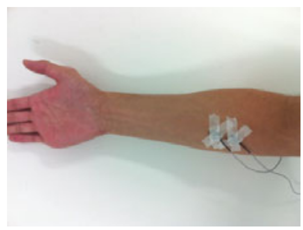


Fig. 6. Flexor carpi radialis muscles

In addition, electromyography (EMG) was performed on the muscles in their forearms and hands to determine the physical characteristics in touch comfort and discomfort. Based on our pilot study, the lumbrical muscles of the hand (Fig. 4), the flexor carpi ulnaris (FCU) muscles (Fig. 5), and the flexor carpi radialis (FCR) muscles (Fig. 6), which remarkably reacted when the sheet switch was touched, were selected as the measured points. To eliminate individual errors and environmental effects, reference data parameters were set prior to the testing trials in the following manner: A digital force gauge indicating the level of force in real time was set at the same height as the model switches, and the participants pressed it with a force of 10 N for 2 s in each vertical and horizontal touch test. The EMG data measured on each of the three points were then used as the reference data. Moreover, the instant when their fingers touched a model switch was recorded using a digital video camera.

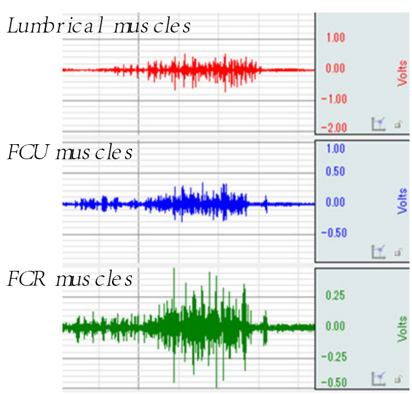


Fig. 7. Example of EMG data

2.6 Experimental Systems

In this experiment, we used an EMG record system (MP150/EMG100C, BIOPAC Systems, Inc.), a digital force gauge (FPG-5, NIDEC-SHIMPO CORPORATION), and a digital video camera (EX-FC100, CASIO).

3 Analysis

3.1 EMG Data

The interval for analysis was set as the time duration from the instant a finger touched the top of the switch to the instant it was withdrawn. Fig. 7 shows an example of the EMG data of each muscle. We calculated the root mean square (RMS) of the EMG data and used the RMS ratio as a measure of muscle activity: RMS ratio = (RMS of trial data)/(RMS of reference data).

3.2 Subjective Scores

The actual scores assigned by the participants were used as subjective scores.

4 Experimental Results

4.1 Results of Vertical Touch Testing

Subjective Evaluation. To examine the degree of comfort elderly users experienced upon touching model switches with different features, we focused on the average score calculated for each switch. Fig. 8 compares the results of the older and younger groups. The figures on each chart show the average scores of each combination of depth and reaction force with diameters 8, 18, and 28 mm. The results indicate that the 8-mm model switch was adjudged uncomfortable by both groups, almost regardless of other parameters. The 18- and 28-mm model switches were adjudged comfortable on average by both groups. However, the model switches that were assigned remarkably high scores were different for both groups. The younger group preferred switches with small depth and reaction force, whereas the older group preferred switches with small reaction force but large depth.

In other words, the younger group experienced comfort when they merely touched soft-feeling switches; whereas, the older group experienced comfort when they touched switches that produced a small "clicking" feel but required deep pushing.

Muscle Activity. To examine the characteristics of muscle activity when participants of the older group touched each switch, we focused on the relationship between the activity level of each muscle and subjective scores. Figs. 9, 10, and 11 show the results of the lumbrical, FCU, and FCR muscles, respectively, so that both groups can be compared. In these charts, the levels of the three parameters are added to the plots corresponding to each model switch. The results indicate that the RMS ratio representing the activity level in the FCU and FCR muscles was higher in the younger group than in the older group; however, the RMS ratio in the lumbrical muscles was higher in the older group than in the younger group. This implies that the older group tended to predominately use hand muscles, whereas the younger group tended to predominately use forearm muscles. Moreover, the younger group tended to assign high scores when the activity level of forearm

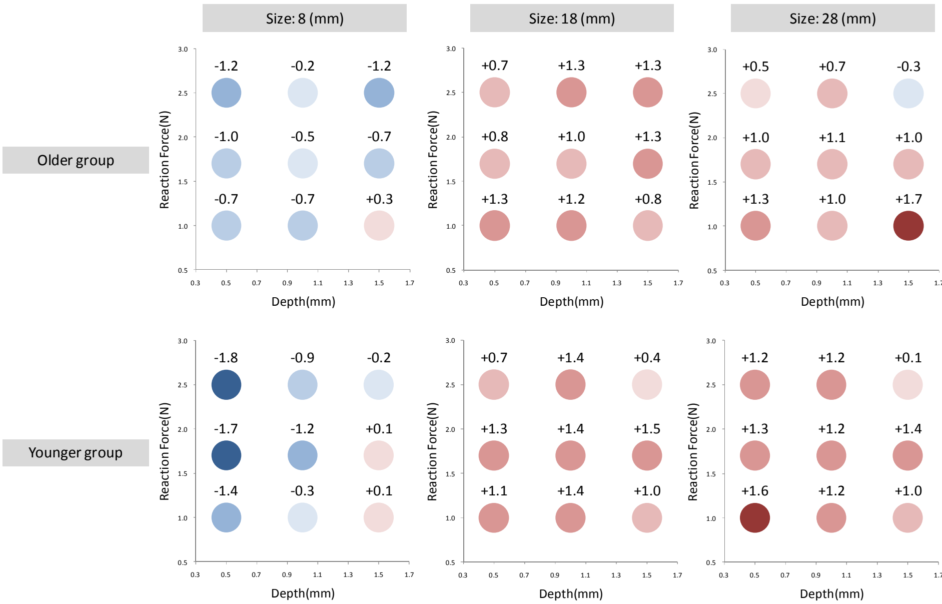


Fig. 8. Average subjective scores assigned by both groups for each model switch in vertical touch testing

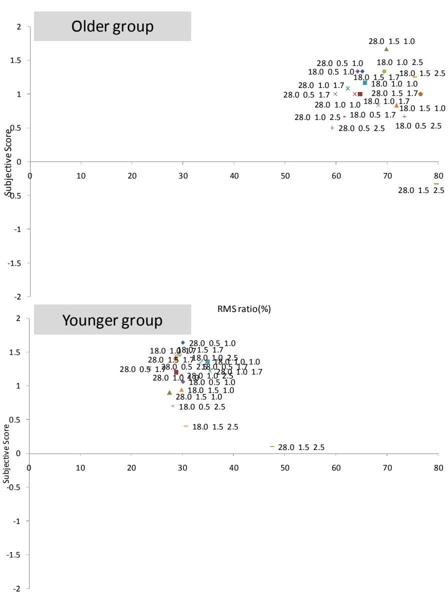


Fig. 9. Relationship between subjective scores and RMS ratio of lumbrical muscles in vertical touch testing

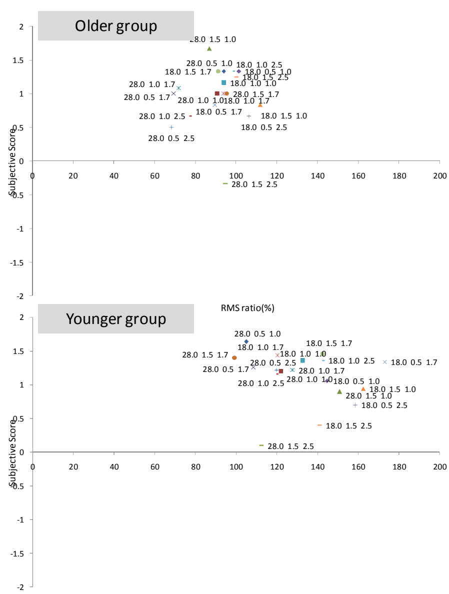


Fig. 10. Relationship between subjective scores and RMS ratio of FCU muscles in vertical touch testing

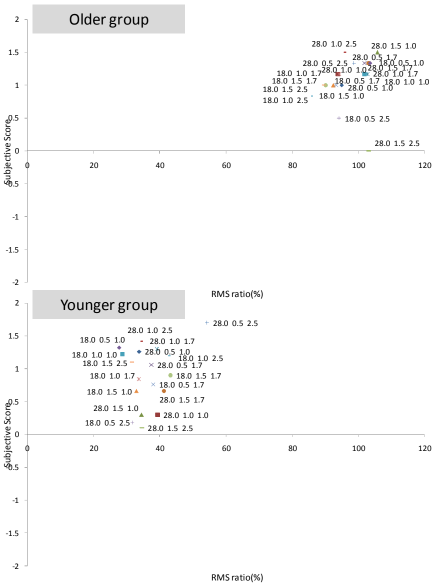


Fig. 11. Relationship between subjective scores and RMS ratio of FCR muscles in vertical touch testing

muscles was smaller, although there is no significant correlation between the activity level of muscles and subjective scores.

4.2 Results of Horizontal Touch Testing

Subjective Evaluation. We analyzed the average score assigned to each model switch in horizontal touch testing in a manner similar to that in vertical touch testing. Fig. 12 compares the results of both groups, who rated the 8-mm model switch as uncomfortable, regardless of other parameters. Moreover, both groups deemed the 18- and 28-mm model switches as comfortable on average. However, the features of each group that were adjudged comfortable showed a discrepancy. The younger group tended to prefer model switches with comparatively small depth and favored model switches with considerable reaction forces.

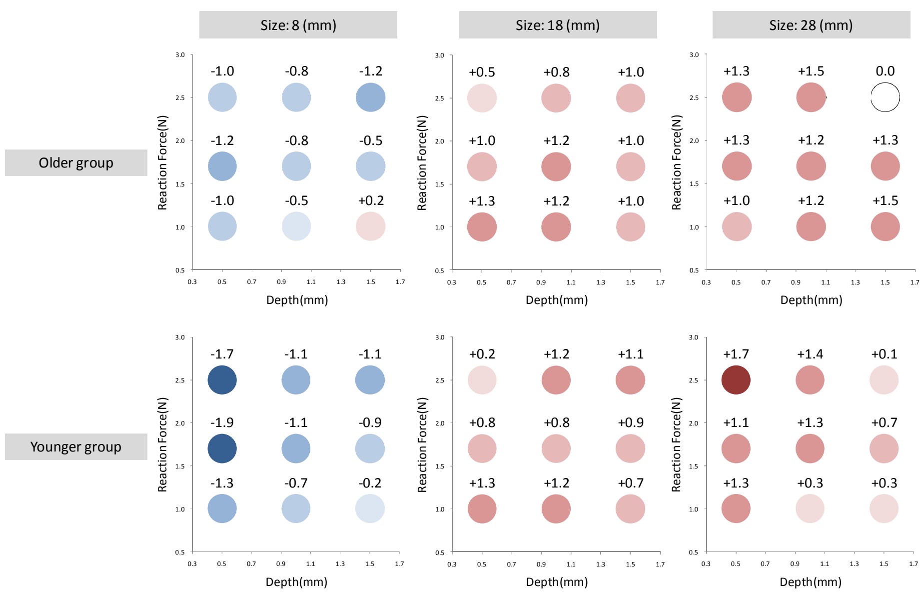


Fig. 12. Average subjective scores assigned by both groups for each model switch in horizontal touch testing

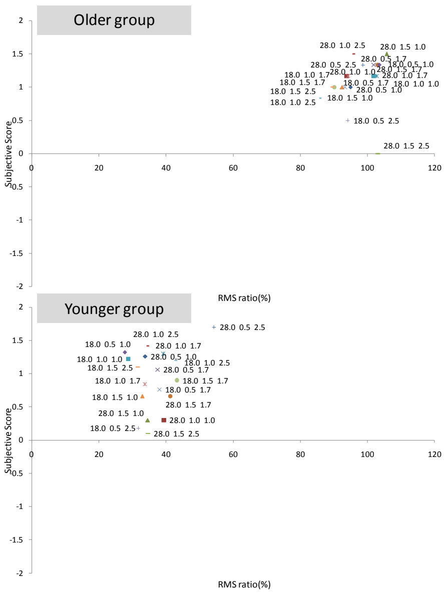


Fig. 13. Relationship between subjective scores and RMS ratio of lumbrical muscles in horizontal touch testing

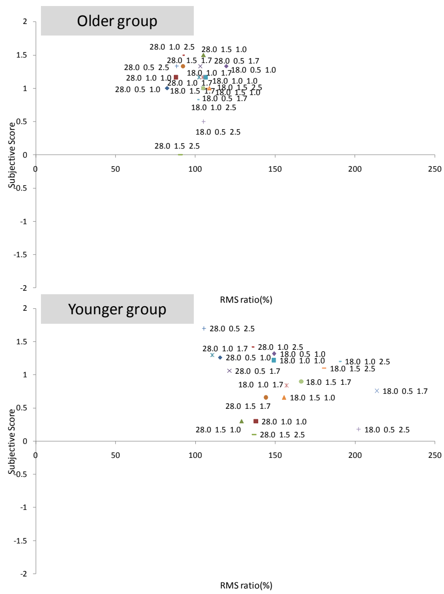


Fig. 14. Relationship between subjective scores and RMS ratio of FCU muscles in horizontal touch testing

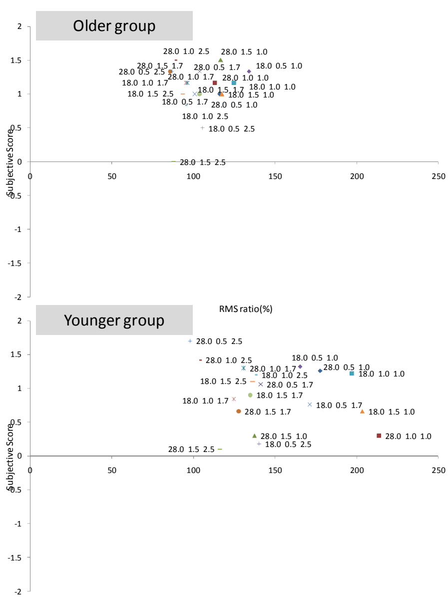


Fig. 15. Relationship between subjective scores and RMS ratio of FCR muscles in horizontal touch testing

Conversely, the older group did not show any remarkable trend during horizontal touch testing.

Muscle Activity. We further analyzed the relationship between the activity level of each muscle and subjective scores. Figs. 13, 14, and 15 compare the results of lumbrical, FCU, and FCR muscles, respectively, between the younger and older groups. The results also indicate that, when touching the model switches, the older group tended to favor the hand over the forearm as compared to the younger group. Moreover, although each RMS ratio of the FCU muscles and of the FCR muscles negatively correlated with the subjective scores in the younger group, each of the three muscles showed little correlation with the subjective scores in the older group.

5 Discussion

We observed that elderly users experienced comfort when touching sheet switches with certain characteristics. In vertical and horizontal touch testing, the 8-mm sheet switches were deemed uncomfortable, and the 18- and 28-mm sheet switches were generally rated comfortable. The standard size of sheet switches presently included in most electrical products is approximately 18 mm; therefore, the standard or a larger size should guarantee comfort for the elderly. Moreover, when the size restriction was satisfied, the comfort experienced on touching sheet switches varied depending on the balance between depth and reaction force. A noteworthy finding is that, in vertical touch testing, older users experienced comfort upon touching sheet switches with a larger depth but requiring a smaller reaction force. On the other hand, younger users experienced comfort upon touching sheet switches with smaller depth and reaction force. In addition, a comparison of their subjective scores with their muscle activity indicates that certain features caused them to experience comfort. In the action of touching a sheet switch, the activity level of forearm muscles was higher in the younger group than in the older group; however, the activity level of hand muscles was higher in the older group than in the younger group. This implies that older users tended to predominately use hand functions to push sheet switches, indicating that they possibly experience comfort to the highest degree when the sheet switch is sufficiently deep to be pushed by moving their fingers.

The difference in the favored muscle activity between the two groups may be attributed to their physiological and cultural backgrounds. At this point, we are particularly interested in the latter effect. That is, we believe that because an increasing number of touch-screen devices has been developed and popularized among the youth, younger users are inclined to touch switches as if operating touch screens rather than hard switches.

6 Conclusion

In this study, we explored the parameters related to the comfort older users experience upon touching sheet switches. Based on the results, we determined that the balance between depth and reaction force is an important factor in designing highly comfortable sheet switches. Moreover, we demonstrated that elderly users tend to predominately use hand muscles to touch sheet switches, as opposed to younger users. Although we cannot deny that the number of participants was not sufficient, our study provided meaningful knowledge to direct our further research.

Acknowledgements. We would like to thank Prof. Sakae Yamamoto and Dr. Masato Sakurai who provided helpful comments and suggestions.

References

1. Nikkei Electronics, 1044, pp. 30–49. Nikkei Business Publications (2010)
2. Research Institute of Human Engineering for Quality Life (HQL), Research study on Construction of a Database for Physical Functions (2000)

3. Fujikake, K., Mukai, M., Kansaku, H., Miyoshi, M., Omori, M., Miyao, M.: Readability for PDAs and LCD monitors among elderly people. Ergonomics 40(4), 218–227 (2004)
4. Mori, M., Horino, S.: Usability enhancement of cellular phone considering elderly people's use. Ergonomics 40(suppl.), 254–255 (2004)
5. Hirano, K., Daimon, T., Toyoda, K.: Preliminary Study on Operational Characteristics of Aged People in Predictive Methods for Japanese Text Input on Mobile Phone. Ergonomics 40(suppl.), 272–273 (2004)
6. Yamagiwa, T., Yoshimura, I.: A Study of Haptic Perception when using a Mobile Device: Argument Focusing on the Influence of Rotation. Japan Industrial Management Association 54(2), 83–94 (2003)
7. Akutsu, M., Hayashi, H., Ikezawa, Y., Ouyama, A., Seino, A., Takeuchi, H.: Agrex INS.: An Experimental Study on Requirements of Computer for the Aged. Ergonomics 43(suppl. 2), 272–273 (2007)
8. Ishihara, K., Nagamuchi, M., Osaki, H., Ishihara, S., Tsuji, A.: Affects of age-related yellowing crystalline lenses on daily life. Ergonomics 43(1), 9–16 (1998)
9. Sakagami, M., Nakashima, Y., Takamatsu, M., Nakajima, S., Mima, K.-i.: Fundamental Research of Visual Barrier -free for Elderly people-On the LED Information Board. Ergonomics 43(suppl. 2), 388–389 (2007)
10. Fujikake, K., Sakurai, T., Miura, K., Hasegawa, S., Takada, H., Omori, M., Matsumura, Y., Honda, R., Miyao, M.: Evaluation of the display quality of LCD for elderly people - Comparative study using Stabilometers. Ergonomics 42(suppl.), 420–421 (2006)
11. Moriwaka, M., Murata, A., Kawamura, M.: Display design in the light of age, task difficulty, character size and background color. Ergonomics 42(2), 132–138 (2006)
12. Kawata, A., Fukumoto, I.: A study of the universal alarms for both young adults and the elderly. Ergonomics 36(5), 261–272 (2000)
13. Kurakata, K., Kubo, Y., Kizuka, T., Kuchinomachi, Y.: Relationship between hearing levels of the elderly and preferred volume levels of television. Ergonomics 34(suppl.), 222–223 (1998)
14. Kuchinomachi, Y.: The characteristics of the elderly people and universal design. Ergonomics 35(suppl. 2), 66–69 (1999)
15. Kurakata, K., Matsushita, K., Kubo, Y., Kuchinomachi, Y.: Auditory signals of electric home appliances (Third report): Temporal ringing pattern. Ergonomics 36(3), 147–153 (2000)
16. Kurakata, K., Matsushita, K., Kuba, Y., Kuchinomachi, Y.: Audio signals in electric home appliances evaluated in terms of the hearing ability of older adults (Second report). Ergonomics 35(4), 277–285 (1999)
17. Kurakata, K., Kuba, Y., Kuchinokahi, Y., Matsushita, K.: Audio signals in electric home appliances evaluated in terms of the hearing ability of older adults. Ergonomics 34(4), 215–222 (1998)
18. Okuda, M., Sato, K.: User Interface Design for Mobile Network Age. Sharp Technical Journal 81, 36–40 (2001)

Support for Generation of Sympathetic Embodied Awareness: Measurement of Hand Contact Improvisation under Load Fluctuation Stress

Takabumi Watanabe[1], Yoshiyuki Miwa[1], Go Naito[1], Norikazu Matsushima[1], and Hiroko Nishi[2]

[1] Waseda University, 3-4-1, Ohkubo, Shinjuku-ku, Tokyo, Japan
[2] Toyo Eiwa University, 32, Miho-cho, Midori-ku, Yokohama, Kanagawa, Japan
takabumi@fuji.waseda.jp, miwa@waseda.jp, go.naito@fuji.waseda.jp, thirdnix-7sphere@suou.waseda.jp, hiroko@toyoeiwa.ac.jp

Abstract. The eventual goal of this study is to establish a method to enhance implicit embodiment and support for creation of expression with others. As a first step, we studied how implicit embodiment relevant to the creation of expressions is represented in the body movement during Hand Contact Improvisation. Specifically, by building a system that measures movements of the entire body in the process of hand contact improvisation, experiments were carried out using skilled person of improvised bodily expressions as subjects. From the results, it was inferred that implicit embodiment is activated and both skilled performers. Next, we have focused on fluctuation as a means to enhance implicit embodiment. Hence, a device was developed to present load fluctuation to the hands of test subjects arbitrarily. Effects of load fluctuation were evaluated using measurement results of the entire body and subjective evaluation. The results suggested that by presenting a fluctuation which has little effect on the consciousness, implicit embodiment of performer can be activated.

Keywords: Bodily expression, fluctuation, embodied interaction, embodiment.

1 Introduction

Because embodiment plays an important role in communication, both participants' bodies need to be on a common actual field [1]. For example, in an expressive body movement in which performers create improvised expressions while keeping their hands in contact with the other's (hand contact improvisation), performers create one bodily expression with a new image, while reading their partner's weight or movement, as well as their mind, through physical contact. In doing so, each performer would feel the sensation as if the body of partner is an extended part of his/her own body, while the partner's movements bring the sensation as if his/her own body has become a part of partner's body. As the relationship between the partners deepens, the inherent framework

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 508–518, 2011.

of individuality, as in "my movement" or "your movement", will be eventually replaced by a new sense of association (sympathetic embodied awareness [2]), in which body movements of respective partners are perceived as "our movements". In other words, there will be a self-other inseparable relationship where one feels the sensation of "embrace the partner while being embraced" or "be embraced while embracing the partner".

The authors aim to establish a media technology that promotes the development of such a relationship and supports the co-creation of expressions. So far the authors have attempted to capture the process of hand contact improvisation to gain basic knowledge [3]. As a result, it was found that implicit embodied interaction may play an important role in the co-creation of expressions by two performers.

This study aims to establish a methodology to enhance such implicit embodiment and support for co-creation expression of between partners. As a first step, the authors decided to study how implicit embodiment relevant to the creation of expressions is represented in the body movement. Specifically, by building a system that measures movements of the entire body in the process of hand contact improvisation, experiments were carried out using skilled person of improvised bodily expressions as subjects. Next, the authors have focused on fluctuation as a means to enhance implicit embodiment. It is generally known that fluctuation elicit implicit information by enhancing the sensitivity of senses [4] [5] and emergence of order [6]. Hence, a device was developed to present load fluctuation to the hands of test subjects arbitrarily. Effects of load fluctuation were evaluated using measurement results of the entire body and subjective evaluation. The evaluation results are provided in the following sections.

2 Measurements of the Entire Body during Hand Contact Improvisation

From earlier studies of the authors regarding measurements of the process of hand contact improvisation, a possibility was found that implicit embodied interaction are made between the partners prior to exchanges are made on the conscious level in a hand contact improvisation [3]. Therefore, in this study, an attempt was made to study how implicit embodiment relevant to the creation of expressions is represented in the body movement.

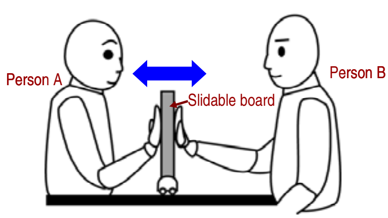


Fig. 1. Slidable-board-mediated bodily interaction

Actual hand contact improvisations performed have a high degree of freedom resulting in a huge number of movements to be measured, which makes the measurement very complex. For this reason, the authors have limited the degree of freedom to "1", which is the back and forth movement, using a mechanism in which a board (slidable board) slides back and forth on a rail as shown Fig.1. Incidentally, it has already been confirmed in past studies [3] that this limitation of movement does not prohibit performers from co-creating expression by sharing an image. Measurements of the entire body movement were done in the process of a hand contact improvisation using the slidable board as an interface. A motion capture system (Optitrack) was used to measure the positions of subjects' back of the head, back, and left elbow. In addition, the position of slidable board was measured by using a linear encoder attached to the board. The subjects (in their forties, female, right-handed) were very skilled in hand contact improvisation as they have long been performing improvisation of bodily expressions using the method. Fig. 2(a) shows the scene of measurement.

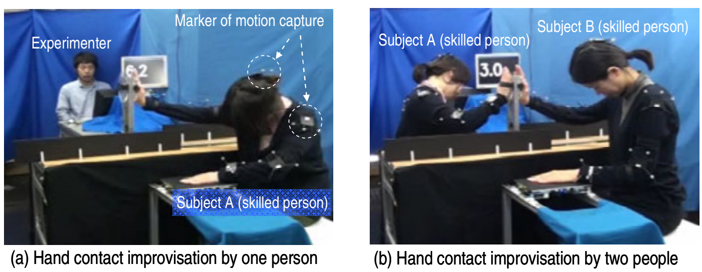


Fig. 2. Scene of measurement of the entire body movement during hand contact improvisation

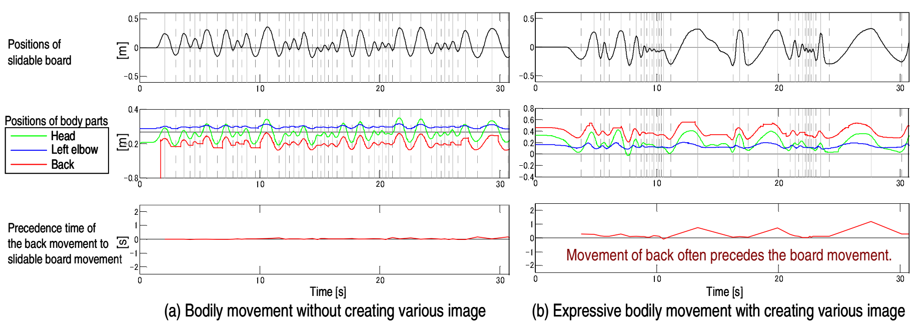


Fig. 3. A part of results of measurement of hand contact improvisation by one person

Measurements were done under two conditions: in one case the subjects were asked to move their hands as a simple exercise without creating a mental image and, in the other, they were asked to use their hands to express the mental image being generated. Measurements were made for the duration of 60 seconds in each case. Fig. 3 shows a part of measurement results. In addition to the positions of slidable board, subjects' back of the head, left elbow, and back, the time the back movement precedes the board movement is shown in Fig. 3. When hand was moved as mere exercise (Fig. 3(a)), it is evident that the board movement is monotonous and that the board movement and the back movement are synchronized. On the other hand, when the subject performed hand contact improvisation (Fig. 3(b)), it was found that the back movement precedes the board movement. This suggests that implicit embodiment, such as a pre-stock motion to express a mental image, may have appeared in the body movements preceding the board movement.

Next, two skilled performers in bodily expressions (in their thirties and forties, female, right-handed) were asked to create expression at their will for 120 seconds. Fig. 2(b) shows the scene of this experiment, while Fig. 4 shows various data regarding movements of the slidable board and subjects' bodies. From the comments of subjects saying that "they felt like there was synchronization at a very deep level from the beginning to end", it is assumed that there was a sensation very close to sympathetic embodied awareness. By focusing the time the body movements preceded the board movements, it is observed that the movements of both subjects often preceded the board movements. From these results, it is inferred that implicit embodiment is activated and both performers.

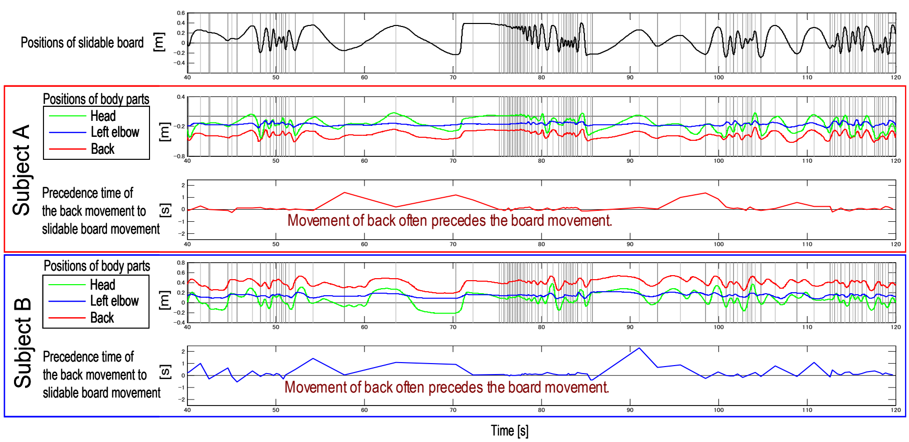


Fig. 4. A part of result of measurement of hand contact improvisation by two people

Based on these measurement results of hand contact improvisation performed by skilled performers, it is assumed that, in order to support co-creation of bodily expression, a new technique should be developed to vitalize implicit embodiment of each person's body.

3 Support for Hand Contact Improvisation Focused on Load Fluctuation

3.1 Concept

Based on the result of Chapter 2, the method which activates implicit embodiment and support for creating bodily expression is considered in this chapter. A phenomenon called stochastic resonance is well known, which becomes tangible and measurable when a noise is applied to a latent weak signal. The stochastic resonance has been testified to exist in a variety of human sensations such as visual, auditory and tactual sensations [4], [5]. It has been found out from a research concerning the neural network of the brain that noise has a role as energy for creating an order, and generation of a self-organizational ordered pattern was investigated [6].

In this way, it is thought that fluctuation has the function which elicits implicit information. Based on the above description, this research is intended to activate implicit embodiment by applying fluctuation as a load during hand contact improvisation. During the hand contact improvisation, a movement of the body of the other person is sensed through a sense of force of the palm. Accordingly, we decided to develop an apparatus for presenting a fluctuation of force to the palm in contact with the palm of the other person as a load. Specifically, a load of the fluctuation is applied to the performer by driving the slidable board as shown Fig.5.

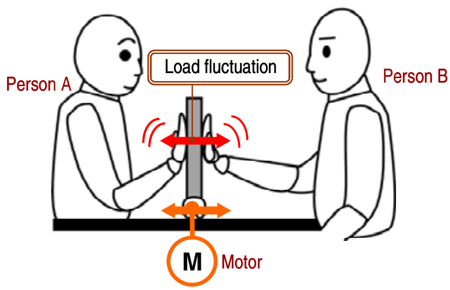


Fig. 5. Hand contact improvisation utilizing load fluctuation display

3.2 Development of Load Fluctuation Display

Fig. 6 shows hardware developed. With the movement of the slidable board restricted to degree of freedom 1 by a linear guide rail (movable range 800 mm), the board (150 x 120 mm) to whose guide is attached is driven with a DC motor to represent a sense of force on the palm. Assuming that a force (about 10 kgf) produced upon a hand contact improvisation is maximum output, an output of the DC motor needs to be amplified using a reducer. However, the back-drivability of the board is reduced considerably by friction and resistance during deceleration by gears, so that when a performer applies an instantaneous external force to it, he or she may feel that the apparatus is heavy. In addition, no smooth and accurate presentation of the sense of force can be achieved due to variance produced between a reaction force generated on the palm and a motor input. Thus, this apparatus utilizes a wire pulley mechanism ensuring a high reverse movability. That is, inertia and friction loss are reduced by a mechanism which winds up a wire attached to both the front and rear sides of the board using a motor (Maxson rating: 170 mNm) via a pulley. Specifically, a following mechanism was invented: two wires are used and one end of each wire is attached to both surfaces of the board with tension applied by a spring and the other end is wound around a threaded shaft (10 mm in diameter), so that the two wires are wound around the shaft inversely. As a result, when one wire is wound up, the other wire is released by an equal length without any relaxation (Fig. 6(b)).

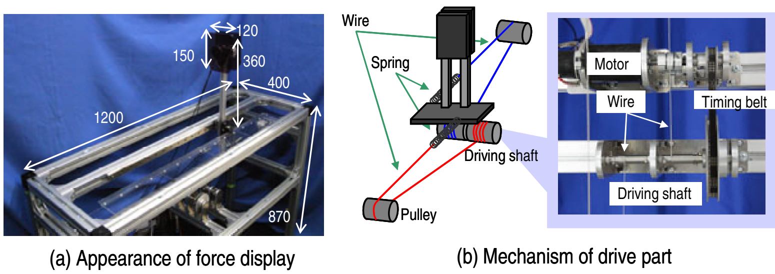


Fig. 6. Hardware of load fluctuation display

The control of this apparatus cancels friction and inertia increased by installation of the drive unit to allow the board to be pushed and pulled smoothly as if no drive unit is installed. After that, fluctuation is applied. For this purpose, it is necessary to know at least loads by friction and inertia accurately. Then, force sensors are attached to the front and rear sides of a frame mounted on a slide mechanism. At this time, a sum of the inertia and frictional resistance of the slidable board are measured as a difference of values from the two sensors. Then, the control was made to calculate friction and inertia and bring them near 0 kgf. For the fluctuation load, fluctuation data is produced preliminarily and added to a target value. The content described above is achieved by feedback PID control using a microcomputer (H8-3048).

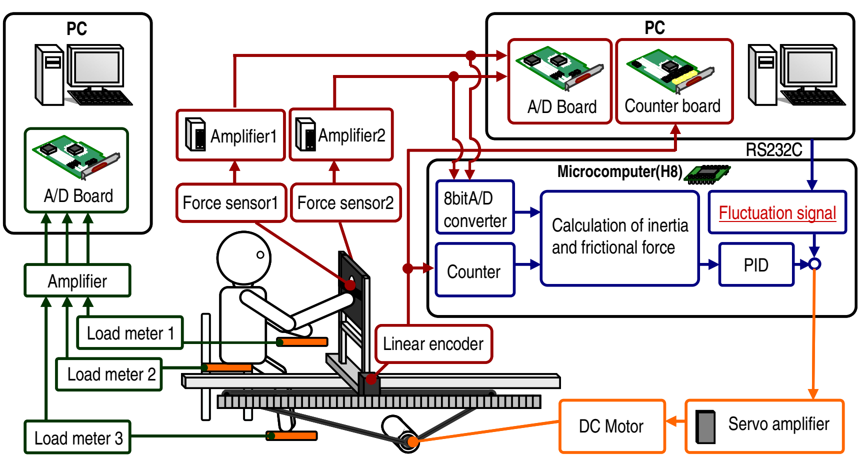


Fig. 7. Composition of the of load fluctuation display system

In addition to the above, the measurement system grasping the motion of the entire body is also explained. As a method for this, changes in position of the center of reaction force from the entire body produced by a bodily movement were measured for this research. Load meters are arranged at all the surfaces which the body contacts, e.g., a table on which the left hand is to be placed, seating face and foot portion in order to calculate a shift of the weight of the body from the weight applied to the load meter and the distribution of the load meters. Nintendo's balance Wii board was used for the load meters at the foot portion and the seating face. To measure the shift of the weight of the left hand, a small load meter in which load cells are arranged on four corners (manufactured by Minebea, rating: 29.42 N) was manufactured and utilized.

The composition of the above whole system is shown in Fig. 7.

4 Load Fluctuation Test for Hand Contact Improvisation

4.1 Load Fluctuation Test for Hand Contact Improvisation by One Person

Firstly, the effect of load fluctuation for body improvisation was to be investigated and this was to be carried out by one person. There were 6 testers (male students in their 20's, right-handed). To begin with, an examination was carried out to find out what kind of impact the size of the load fluctuation would have on hand contact improvisation. For the fluctuation in white noise frequency, three types of amplitudes: I: none, II: weak (max amplitude 0.2 [kgf]) and III: strong (max amplitude 1.2 [kgf]) were represented during the hand contact improvisation. The experiment time was 60 seconds for each condition, and the experiment sequence was carried out randomly. After each condition is concluded, a questionnaire was handed out. This consisted of 7 rating system against items shown in table Fig.8.

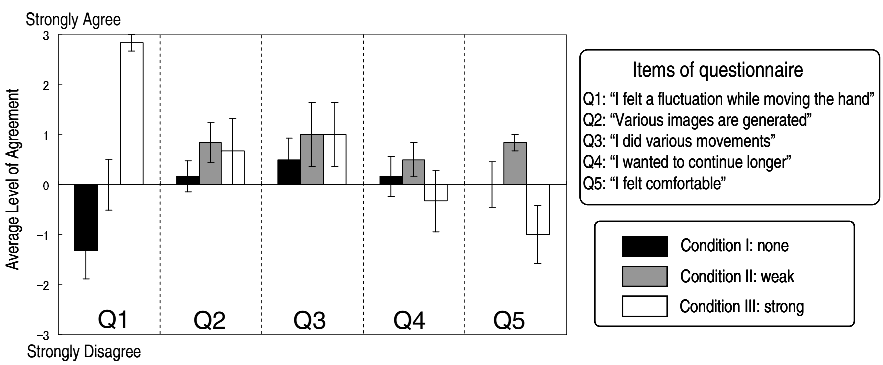
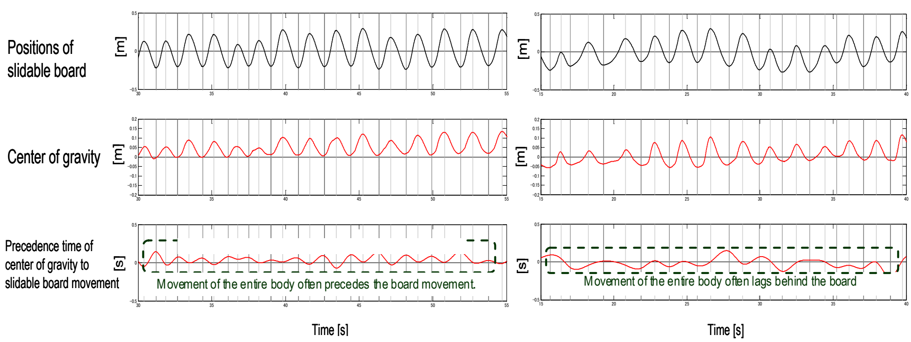


Fig. 8. Result of questionnaire on hand contact improvisation by one person under load fluctuation

Figure 8 shows a graph of the average values for subjective report evaluation from the questionnaire. In Q1: "I felt a fluctuation while moving the hand", condition I and II showed negative evaluation and condition III showed positive evaluation. Hence, it is thought that the fluctuation in condition III has risen strongly. For questions Q2: “Various images started to flow” and Q3: “I did various movements”, the differences in testers were largely scattered, however higher ratings were obtained in comparison to condition I where there were no fluctuations. On the other hand, for questions Q4: “I wanted to continue longer” and Q5: “I felt comfortable”, condition III showed a tendency to fall.

Here, condition II was given weaker fluctuations, which do not affect the consciousness, and condition III was given stronger fluctuations which have an effect on the consciousness. In order to examine the detailed differences between these two conditions, focus was on the measured results of the hand (slidable board) movement and overall body movement. Figure 9 (a) and (b) shows part of the results from one of the tester using the two conditions.

It was understood that the preceded time for the body movement was relatively longer than the hand movement in condition II. On the other hand, the preceded time for the body movement was decreased in condition III, and the body movement appeared to show delay in comparison to the hand movement. In Chapter 2, it has been shown that the body function has a potential to be presented as a pre-stock movement for creating such image of improvisation. Therefore, as for this result, the implicit embodiment was increased when given weaker fluctuations and it has encouraged the creation of bodily-expression. On the other hand, there were also comments such as “I was fixed on the decided image”, when given stronger fluctuations which have an effect on the consciousness. The implicit embodiment was suppressed, and it has interfered with the proactive creation of improvisation.



(a) Condition II: Weak (b) Condition III: Strong

Fig. 9. A part of results of measurements of hand contact Improvisation by two people under t load fluctuation

4.2 Load Fluctuation Test for Hand Contact Improvisation by Two People

Next, tests were carried out where load fluctuation was presented to palms of the subjects during hand contact improvisation performed by two people. The subjects consisted of three pairs of 6 (male students in their 20's, right-handed). In this study, three types of amplitudes: I: none, II: weak (max amplitude 0.4 [kgf]) and III: strong (max amplitude 1.2 [kgf]) were presented for the fluctuation in white noise frequency during the hand contact improvisation. The experiment time was 60 seconds for each condition, and the experiment sequence was carried out randomly in pairs. After all the conditions were concluded, a questionnaire was handed out. This consisted of 7 rating system against items shown in Fig.10.

Figure 10 shows a graph of the average values for subjective report evaluation from the questionnaire. Looking at the evaluation for Q1-Q5, results showed the same tendency with the results of 4.1. For the question items Q6: "I felt unified with the partner's body" and "It was possible to create an expression with the partner", ratings for condition II was the highest, and ratings for condition III tended to be lower. It was suggested that even though the fluctuation has little effect on the consciousness, there was a potential to strengthen the connection between the partners.

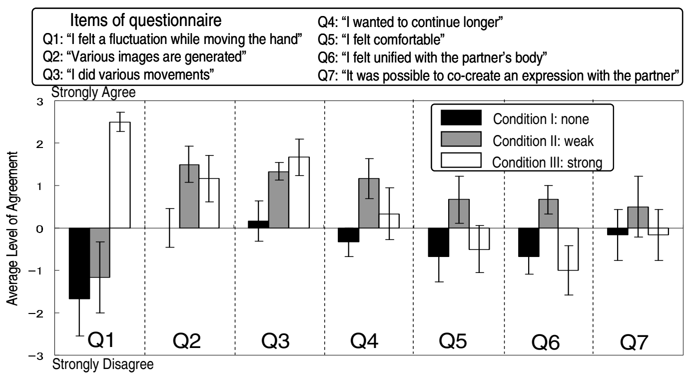
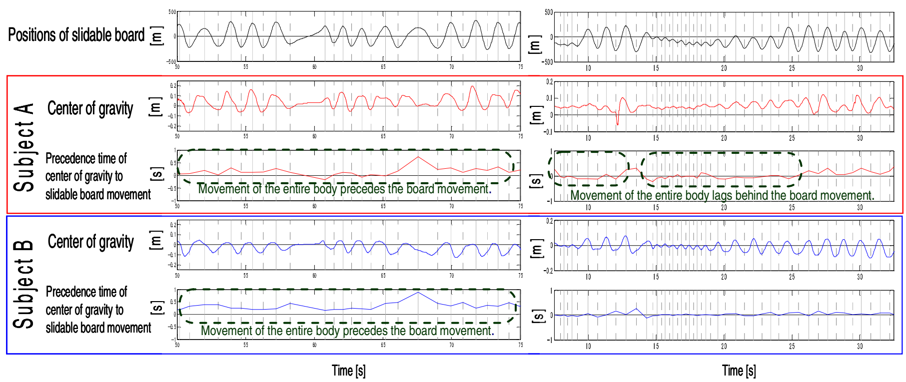


Fig. 10. Result of questionnaire on hand contact improvisation by two people under load fluctuation

Here, the differences in physical interactions were examined in the case where weak fluctuation was provided, which has little effect on the consciousness, and also in the case where strong fluctuation was provided, which has an effect on the consciousness. Figure 11 shows part of the results from a pair of testers for their hand and body movement. Looking at the results for condition II within Fig.11, it is understood that the preceded time for the change in center of gravity of the testers, A and B, were longer against the movement of the hands. Therefore, the implicit embodiment increases and it has encouraged the creation of improvisation as seen in 4.1.

The above results showed that by presenting a fluctuation which has little effect on the consciousness, implicit embodiment of each other can be activated, and we can speculate implicit connection for co-creating expression. The verification and dynamics of this will be in the future study.



(a) Condition II: Weak (b) Condition III: Strong

Fig. 11. A part of results of measurements of Hand Contact Improvisation by two people under two load fluctuation conditions

6 Conclusion

The eventual goal of this study is to establish a method to enhance implicit embodiment and support for co-creation of expression with others. As a first step, we studied how implicit embodiment relevant to the creation of expressions is represented in the body movement during Hand Contact Improvisation. Specifically, by building a system that measures movements of the entire body in the process of hand contact improvisation, experiments were carried out using skilled person of improvised bodily expressions as subjects. From the results, it was inferred that implicit embodiment is activated and both skilled performers. Next, we have focused on fluctuation as a means to enhance implicit embodiment. Hence, a device was developed to present load fluctuation to the hands of test subjects arbitrarily. Effects of load fluctuation were evaluated using measurement results of the entire body and subjective evaluation. The results suggested that by presenting a fluctuation which has little effect on the consciousness, implicit embodiment of performer can be activated.

Acknowledgment. This study was partially supported by a Grant-in-Aid for Scientific Research, 21650048 and 22700221, from the Japanese Ministry of Education, Science, Sports and Culture. This study was also conducted under the project “Generation and Control Technology of Human-Entrained Embodied Media” and is supported by CREST of JST, Project Research "Principal of emergence for empathetic “Ba” and its applicability to communication technology" by RISE Waseda Univ., and the GCOE program “Global Robot Academia.”

References

1. Shimizu, H., Kume, T., Miwa, Y., Miyake, Y.: Ba and co-creation. NTT Publishing (2000) (in Japanese)
2. Nishi, H.: Interactive Communication through body, Doctoral thesis, Kobe University (2003) (in Japanese)

3. Watanabe, T., Matsushima, N., Nishi, H., Miwa, Y.: Electromyography Focused on Activeness and Passiveness in Embodied Interaction: Toward a Novel Interface for Co-creating Expressive Body Movement. Journal of Advanced Mechanical Design, System, and Manufacturing 5(1), 35–44 (2011)
4. Simonotto, E., Riani, M., Seife, C., Roberts, M., Twitty, J., Moss, F.: Visual Perception of Stochastic Resonance. Physical Review Letters 78, 1186–1189 (1997)
5. Miyaoka, T., Mano, T., Ohka, M.: Mechanisms of fine-surface-texture discrimination in human tactile sensation. The Journal of the Acoustical Society of America 105(4), 2485–2492 (1999)
6. Usami, A., Matsuki, N., Ikegaya, Y.: Spontaneous Plasticity of Multineuronal Activity Patterns in Activated Hippocampal Networks. Neural Plasticity 108969 (2008)

Part VI

Adaptation and Personalisation

Different People Different Styles: Impact of Personality Style in Web Sites Credibility Judgement

Rahayu Ahmad[1], Jieyu Wang[1], Karoly Hercegfi[2], and Anita Komlodi[1]

[1] Department of Information Systems, University of Maryland Baltimore County
[2] Budapest University of Technology and Economics, Hungary
{arahayu1,wajieyu1,komlodi}@umbc.edu,
Hercegfi@erg.bme.hu

Abstract. Reliance on the Internet as a source of information has increased dramatically in recent years among information seekers. The reliability of information on the Internet can sometimes be questionable due to the absence of an editorial function. Users need to carefully consider the quality of the information before using it. The goal of this research was to study the process of credibility evaluation by users. We examined how personality styles influence the way people make credibility judgments when they are browsing online. Our results show that thinkers' evaluation style were more fact based than intuitive users who relied more on their initial impression and prior experience with the websites. Perceivers gathered information from various sources before trusting information and forming perception on the credibility of authors and websites. Lastly the judgers tend to formed conclusions quickly. This resulted to filtering credible web sites based on the format and layout of the websites.

1 Introduction

Reliance on the Internet as a source of information has increased dramatically in recent years among information seekers. On the other hand, the advent of Web technologies has created unprecedented opportunity for information producers to easily publish information [6]. The privilege of writing anonymously and the absence of an editorial function can lead people to often questionable or inaccurate information. These processes have created a need for examining the process of credibility evaluation by users. The credibility evaluation styles, however, might vary depending on the skills, experiences and learning style of the user [9]. This study attempts to have a more in depth look at the effect of one particular individual difference, personality, on performing credibility evaluations of Web pages.

2 Background

Credibility is not a property of an object but a quality or perception ascribed to an object by the receiver of the information [9][10]. There are two components of credibility; trustworthiness and expertise. Trust is defined by the terms "well-intentioned, truthful, unbiased" [8, pp.1] meanwhile expertise, is defined by terms such as "knowledgeable, experienced, and competent"[8, pp.1]. Research in this area

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 521–527, 2011.

have identified various factors that can influence credibility judgement. Among the important factors taken into consideration in making credibility judgement are look and design of the page and ease of use. Web sites that look professional and that are easy to navigate are perceived as more credible by users[11]. Other tracks of research in this area examined the role of structural and message features in influencing the credibility evaluation by users [12][13]. Several structural features have been recognized to influence the perceptions of credibility. These include top-level domain name, navigation tools, presence of privacy policy statements, third-party endorsements, site ownership and site contact information [12] [13]. Message features like inclusion of quotes, testimonials, statistics disclosure of authorship, references, information currency (posted dates of content modification), have also been found to contribute to credibility perceptions [12] [13] [14].

Based on a number of large credibility studies, a theory called Prominence Interpretation Theory was later proposed[9]. In this theory Fogg argued that only the features noticed (prominence) by the users will be considered in making credibility evaluation. Many factors influence how these features will be noticed or perceived. Users' ability to scrutinize the information, the topic of information and the users' level of expertise can influence how the features of web sites are perceived. In addition, individual differences for example cultural background, cognitive and personality style may also influence how the features of web sites are perceived and factored into the credibility evaluation process.

Currently, studies that examined the impact of individual differences are still limited. Some researchers have examined the impact of expertice level towards credibility perception of web sites. Generally the findings suggested that novice users relied mostly on the visual and information design web sites and not on other features [1][14]. They were generally satisfied and comfortable when there was a lot of information presented in the web sites without much concern about the sources of information [1]. On the contrary, experts were more comprehensive in their credibility evaluation process. Besides the look and design of the page, they also focused on the name, affiliation and reputation of the Web Sites in judging the credibility of the site [1] [14]. These studies suggested that individual differences play an important role in influencing credibility evaluation styles. Thus our study attempts to examine how user characteristics, particularly their personality style, impact credibility judgments.

3 Methodology

Fifteen people have participated in this study. 70% of them were female while 30% were male. 86.7% were undergraduate students while the rest were graduate students. On average, participants spent 24.6 hours/week searching information from the Web.

Prior to the experiements, participants were required to fill out online questionnaires on their demorgaphic information, cultural background, computer and Web experience, and personality style. We adopted the Champagne and Hogan [4] personality style framework to classify the participants into four types of personalities: intuitive, judging, thinking, and perceiving.

The participants were asked to carry out several information seeking tasks. Three tasks were prescribed for them and one was defined in collaboration with the participants to reflect a more realistic task. The first task was a known-item information-seeking task. This task served as a warm-up period for the participants and the answer was on purpose easy to find. The next two tasks were medium-complexity tasks, one in English and one in Spanish. Both concerned travel planning to a given destination for a certain amount for time. The third task was based on a real task of the participant and was also a topic-driven task. While the participants were searching, their activities were recorded.

For each task, participants were required to select five websites as relevant and bookmark these. Out of these, they were asked to select one that they perceived as the most credible and they rated the credibility components like believability, accuracy, trustworthiness, bias, and completeness of the information in a questionnaire [2] [7]. All items were measured on a five-point scale. Participants also rated the web sites along several dimensions that have been shown to impact credibility such aesthetic appeal, organization of information, timeliness of information, validity of information, matching information needs, matching prior expectations, literary appeal, and profit purpose appeal [10][11][12][13][14].

At the end of the session the participants were interviewed about their general credibility evaluation process. They were asked about the frequency, methods, and criteria of their web credibility evaluation behavior.

4 Analysis

Based on the participants' responses on the personality questionnaires, we identified their strongest personality dimension. Based on their definite strength scores, we then classified them into five different personality groups; Judging (N= 4), Thinking (N= 3), Perceiving (N= 2), Intuitive (N=3) and Sensing (N= 1). We had to omit cases where there were no definite strengths identifiable based on their personality scores. Using the interview data, we performed open and axial coding [5] to identify the emergent trends and styles that participants used in performing credibility evaluation. This analysis is performed within similar personality groups and between different personality groups.

5 Results and Discussion

5.1 Thinking Person

Thinkers make judgments about life, people, occurrences, and things based on logic, analysis, and evidence, avoiding the irrationality of making decisions based on feelings and values [4]. Their personalities were reflected in their credibility evaluation styles. Evaluating credibility of web sites was something natural for the thinkers. They seemed to evaluate the credibility of websites almost every time they went to a site as described by one of our participants:

> *Every time I find something, it will subconsciously go to my mind that it's a credible website or not because I know what I look for, that makes something credible.* *(p7)*

Similarly, when another thinker was asked how frequently they performed credibility evaluation, one of them said:

> *I would say every time I looked at something* (p11)

The thinkers also emphasized their verification of sources of information including the author competencies, as described by another participant:

> *They [the author] could be someone graduated from Stanford and CEO of this company and I said "ohh this could be credible", versus somebody who doesn't have a resume or I am not finding any searches about a particular person when they wrote this article.* (p14)

Our findings demonstrated that thinkers, reflective of their personality, often subconciously think about credibility when searching information from the websites. Their credibility evaluation style were also more fact based, involving verification of authors' background andexpertise, citations and reputation of the information sources.

5.2 Perceiver Person

The perceiver is a gatherer, always wanting to know more before deciding, holding off decision and judgments [4]. The following excerpt vividly demonstrates how an intutive participant still considered information from a potentially unreliable source, and used the information to find other credible pages that match her information needs:

> *Well first I went to Wikipedia to get an idea for what to look for Wikiepedia is not reliable all the time but it will give me an idea for what I could be looking for, and then looking at that [the Wikipedia] gives attractions that Sucre has, I went to Google and typed some stuffs.* (p13)

As the participants in this group were more inclined to gather the whole perspetives, their evaluation style involved navigation from links to another links, helped them in judging the credibility of authors and information on the Wb sites:

> *I trust the links (references) from Wikipedia. Umm like a in the case of one researcher. He has links to his actual work. I did clicked the links to his articles, journals and it was credible journals.* (p5)

Other traits of perceiver are open, flexible, adaptive, nonjudgmental, able to see and appreciate all sides of issues, always welcoming new perspectives and new information about issues [3] as quoted:

> *I usually turn to look at the websites outside the country from different perspectives* (p1)

In performing credibility evaluations, participants belonged to this personality group, were not jumping into conclusion when noticing controversial information. They would verify the information with other sources. The excerpt below shows how our participant verified the information before reaching into conclusions:

> *Well in the Wikipedia article said that he (the author) is under investigation for his violations of certain ethics. So I look more information about him.* *(p5)*

Unlike the thinkers, perceivers were more tolerable to users' reviews and feedback allowing them to get information from various perscpectives before trusting the information on vendors'websites:

> *So I usually click on what's the most popular and what people usually buy. I often read the feedback, the problems they have, I try to look it up, too. What problem I would encounter too with the laptop or computer…I like to look up everything. I don't like to buy something without knowing.* *(p1)*

Overall these findings suggested with respect to their personality style, participants in this group liked to have whole perspectives and considered various factors before believing the websites. They used information from different resources albeit some of them may not be reliable, and later evaluate those information to form their own credibility judgement.

5.3 Intuitive Person

The intuitive person thinks and discusses in spontaneous leaps of intuition that may leave out or neglect details [4]. Our observation suggested that the intuitive person relied much on the initial impression of the websites in making credibility judgment. The participant below explained why the web site is considered believable to her:

> *As far as my first impression is I see it is colorful, catches your eye with moving pictures.... I would say in general to helps draw my attention. The bright colors kind of make me want to read what it says.* *(p12)*

The other participant also emphasized the initial impression in believing the web site*:*

> *It seems to be a very well organized, well-maintained website.* *(p8)*

At the same time, they tended to rely on their familiarity of the web sites in making credibility judgment as demonstrated in the following quotes:

> *Any time I'm looking for information, and I find it on the sites that I'm not familiar with, I usually don't trust it much. I am more comfortable to go to the sites that I'm familiar with.* *(p8)*

> *I used one that I am familiar and comfortable using and I know that it helped me to get from point a to point B before. So I know that it's legitimate, it's credible* *(p12)*

They also used their experience or prior knowledge in judging the websites:

> *If I've heard of it, but never use it before, I'm more likely to trust and use it a lot.* *(p8)*

The intuitive person becomes bored with nitty-gritty details, concrete, actual, and facts unrelated to concepts [4]. Unlike the perceivers, who considered the information from various perspectives, intuitive person paid less attention to supplemental information like user reviews and testimonials:

I learned that I have never really read a lot reviews because it's a matter of opinion. *(p12)*

These findings demonstrated that the intuitive evaluation styles were less fact based compared to the thinkers. Instead of focusing on detailed information, they relied mainly on their 'good feeling' of the web sites and their prior positive experiences with the web sites in making credibility judgement.

5.4 Judging Person

The judger is decisive, firm, and sure, setting goals and sticking to them [4]. People who had more orientation towards a judging personality are likely to come to conclusions quickly [3].This behavior was reflected in their credibility evaluation as demonstrated below:

If it doesn't look professional, I immediately cut it off. *(p2)*

If it is a blog especially if it is something controversial than I don't usually take them as being credible. *(p6)*

I had troubled with it [the website] and I like to do things quickly so I would just give up on it and try another web site. *(p6)*

As judging person preferred structure and organization[3]; they placed great emphasis on the format and layout of the web sites in making credibility judgement. The excerpt below demonstrated the way they formed credibility judgement:

Ummm, it's poorly designed and you can tell there is barely any effort put in, usually if it's an organization they want to prove that you know something, especially if it is a big organization they want to put in something nice, usually the less aesthetically pleasing it is the less I'm likely to go to it. *(p6)*

The presentation is the first to look at. If I go to a website, the content is not well printed out, I do not really trust it. *(p2)*

Probably the first thing I want to stake out is the format. *(p7)*

6 Conclusion and Future Work

This study extends our understanding of the process of credibility judgments. We found personality styles to influence the process of credibility judgment and the recognition of features taken into consideration while making credibility judgments. We have demonstrated how thinkers' evaluation style was more fact based than intuitive participants' who relied more on their initial impression and prior experience with the websites. The perceivers gathered information from various sources before believing and forming a perception on the credibility of authors and websites. Lastly, the judgers tended to form conclusions quickly. This resulted in filtering credible web sites based on the format and layout of the websites. This study is still in progress and further data collection and analysis will be conducted.

References

1. Ahmad, R., Komlodi, A., Wang, J., Hercegfi, K.: The Impact of User Experience Levels on Web Credibility Judgment. In: Proceedings of American Society for Information Science and Technology, Pittsburgh (2010)
2. Austin, E.W., Dong, Q.W.: Source vs content effects on judgments of news believability. Journalism Quarterly 71, 973–983 (1994)
3. Briggs-Myers, I.: Myers-Briggs type indicator Palo A lot, 6th edn. Consulting Psychologists Press, CA (1999)
4. Champagne, D.W., Hogan, R.C.: Supervisory and Management Skills: A Competency Based Training Program for Middle Managers of Educational Systems (1979)
5. Corbin, J., Strauss, A.: Grounded theory research: Procedures, canons and evaluative criteria. Qualitative Sociology 13, 3–21 (1990)
6. Flanagin, A.J., Metzger, M.J.: The role of site features, user attributes, and information verification behaviors on the perceived credibility of Web-based information. New Media & Society 9, 319–342 (2007)
7. Flanagin, A.J., Metzger, M.J.: Perceptions of Internet information credibility. Journalism and Mass Comm. Quarterly 77, 515–540 (2000)
8. Fogg, B.J., Tseng, H.: The Elements of Computer Credibility. In: Proceedings: Conference on Human Factors and Computing Systems (CHI), pp. 80–87 (1999)
9. Fogg, B.J.: Prominence-Interpretation Theory: Explaining how people assess credibility online. In: Proceedings: Conference on Human Factors in Computing Systems (CHI), pp. 722–723 (2003)
10. Fogg, B.J., et al.: What Makes Web Sites Credible? A Report on a Large Quantitative Study. In: Proceedings: Conference on Human Factors in Computing Systems (CHI), pp. 61–68 (2001)
11. Fogg, B.J., Soohoo, C., Danielson, D.R., Marable, L., Stanford, J., Trauber, E.R.: How do users evaluate the credibility of Web sites? A study with over 2,500 participants. In: Conference on Design for User Experience (2003)
12. Hong, T.: The influence of structural and message features on Web site credibility. Journal of the American Society for Information Science and Technology 57, 114–127 (2006)
13. Rains, S.A., Karmikel, C.D.: Health information-seeking and perceptions of website credibility: Examining Web-use orientation, message characteristics, and structural features of websites. Computers in Human Behavior 25, 544–553 (2009)
14. Rieh, S.Y.: Judgment of information quality and cognitive authority in the Web. Journal of the American Society for Information Science and Technology 53(2), 145–161 (2002)
15. Stanford, J., Tauber, E., Fogg, B.J., Marable, L.: Experts vs. Online Consumers: A Comparative Credibility Study of Health and Finance Web Sites (2002), http://www.consumerwebwatch.org/dynamic/web-credibility-reports-experts-vs-online-abstract.cfm

A Comprehensive Reference Model for Personalized Recommender Systems

Matthias Breyer, Kawa Nazemi, Christian Stab, Dirk Burkhardt, and Arjan Kuijper

Fraunhofer Institute for Computer Graphics Research,
Fraunhofer Str. 5, 64283 Darmstadt, Germany
{matthias.breyer,kawa.nazemi,christian.stab,dirk.burkhardt,
arjan.kuijper}@igd.fraunhofer.de

Abstract. Existing reference models for recommender systems are on an abstract level of detail or do not point out the processes and transitions of recommendation systems. However, this information is relevant for developers to design or improve recommendation systems. Even so, users need some background information of the calculation process to understand the process and accept or configure these systems proper. In this paper we present a comprehensive reference model for recommender systems which conjuncts the recommendation processes on an adequate level of detail. To achieve this, the processes of content-based and collaboration-based systems are merged and extended by the transitions and phases of hybrid systems. Furthermore, the algorithms which can be applied in the phases of the model are examined to identify the data flow between these phases. With our model those information of the recommendation calculation process can be identified, which encourages the traceability and thus the acceptance of recommendations.

Keywords: recommender system, recommender reference model, recommendation calculation, calculation information, data flow.

1 Introduction

In internet-based information systems the user is confronted with a growing amount of data, which makes it difficult to identify the relevant information. This is not only a problem for information seekers but also for information providers. In particular for commercial enterprises selling products on the internet, the demand for offering adequate information to a user is crucial. The financial success of their online selling platform directly depends on the user-centered identification and presentation of relevant information [10].

To assist the user in finding relevant information, three complimentary methods can be applied: classification and structuring, search-engines, or recommendation systems. The classification and structuring approach is suitable for an explorative system in which the user can browse through and easily gather an overview of the information. This may be time consuming and ineffective for non-explorative usage scenarios. Search engines are an established approach commonly users are familiar with. The main drawback of search engines is the need for appropriate search queries

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 528–537, 2011.

the user has to declare. Especially if the user has no clue about available information this problem deteriorates. To overcome these issues recommender systems (RS) can be integrated to identify and present information artifacts to the users which match their individual needs and preferences. Using RS the user can benefit from these recommendations with a minimum of inputs specifically stated to the RS. Furthermore, recommendation systems as merchandising tools individually advertise products and thus can enable increasing sales.

In this paper we present a comprehensive reference model for personalized recommendation systems. The model unifies relevant aspects of existing and established PRS reference models. The objective of this reference model is to show the overall interplay of the components and coevally accenting the impact of input parameters and transferred data in the recommendation calculation process. As far as we know there exists no such reference model on this level of detail.

The resulting PSR model is a connected and comprehensive reference model for existing personalized recommendation approaches comprising relevant detail information. The comprehensive reference model is especially suitable for further developments and optimizations of PRS.

In the next section we introduce basic information for recommender systems. The third section will give an overview about related works on which the presented reference model for recommender systems is built on. The comprehensive reference model itself will be described in section four. The last section summarizes the paper.

2 Basics for Recommender Systems

For a common understanding of the construction of a comprehensive reference model for recommender systems some basic information is essential. In this section these basics will be described.

2.1 User Feedback as Input Data

To obtain user and usage data as input data for recommender systems basically three approaches exists which can be applied in combination [8]:

Explicit feedback/direct learning technique. The user explicitly states preference or socio-demographic information in input fields of the 'conversational system'. The main drawback of this approach is users have to be motivated to actively perform these inputs [10]. Furthermore, users have to interpret and understand the feedback mechanisms correct in order to achieve proper recommendations [2].

System-driven explicit feedback/partially direct learning technique. The user explicitly states preferences of socio-demographic information in input fields, but the system determines when and which information is necessary. The objective is to minimize the number of explicitly stated information thus minimizing user interactions which are not related to his originally task.

Implicit feedback/indirect or transparent learning technique. User preferences are deduced by analyzing the interactions ('behavioral usage data', like mouse movements, selections, orders, etc.) of the user [18].

2.2 Classification Criteria: Personalization and Functional Input Sources

Recommender systems are mainly classified in personalized, non-personalized, and supporting systems [1]. In between personalized and non-personalized systems can be settled a fourth category named group-based recommendation systems [7]. In this paper we focus on personalized recommender systems (PRS) which predict individual preference values for information artifacts [13].

For PRS various classification criteria were identified in previous works, of which the criterion by functional input seems to be used most often. To support comparability this criterion will be used in this paper equally. This criterion partitions PRS in content-based, collaboration-based, and hybrid approaches [1,11]. Content-based PRS (item-based/feature-based) extract significant features out of the content (domain-space), based on which similar information artifacts are identified to these artifacts the user rated positive (implicit or explicit, see 2.1). In contrast, collaboration-based RPS compare characteristics from the user models itself to recommend information artifacts ranked positive by similar user models (targeted customer inputs and/or community inputs). Hybrid PRS combine both of these approaches to overcome drawbacks of each separately.

With the evolving Web2.0 and social communities the class of group-based PRS (social-based, social-tagging-based) is stated as a fourth category [10]. Due to the fact these systems use similar functional input sources these systems will here be classified as collaboration based systems. Furthermore, there exist the categorizations of rule-based and knowledge-based systems. However, these require predefined rules and knowledge, so they are not fully automatic and will not be considered further on [5].

3 Related Work

After this short introduction to the basics for recommendation systems and the classification criteria applied in this paper, in this section related works will be discussed. These related works are also the basics for the designed comprehensive reference model in the subsequent section.

3.1 Process Models for Recommender Systems

In the field of recommendation systems some models exist which describe the process of recommendation calculation in different abstraction levels. Often, these models are influenced by the point of view the model is developed for. To obtain an overview of these models the most important are presented here, which lay the groundwork for the comprehensive reference model.

Abstract model of recommendation systems. [5] described this model which consists of four components: (i) background knowledge: knowledge domain and previously gathered user/usage data (see section 2.1), (ii) user/usage data of the current interaction session, (iii) an recommendation algorithm which calculated user preference approximations based on the background knowledge and the current user/usage data, and (iv) presentation of the recommendations.

Model including activities and information flows. [18] present an extended recommendation model which includes activities between the involved responsibility parties and the resulting information flows between these parties. This model is defined on a very abstract level of detail to cover as many recommendation activities as possible. For a concrete usage scenario these activities significantly vary. Furthermore, not each of the activities of information flows has to be concretized or realized.

Conceptional classification model. A particularized description model is presented in [15]. In this model the characteristics of recommendation systems are differentiated in functional in- and outputs (see section 2.2), recommendation method (see section 3.3), and additional design decisions. The model differentiates the functional inputs in those that describe the current user and those representing the whole community. Additional design decisions are determined by the recommendation system provider (before runtime) but determine the recommendation quality and the kind of presentation. Thus, these decisions have a major impact for the users' perception of the recommendations which furthermore constitute the success of recommendation systems.

3.2 Recommendation Calculation Processes for Recommender Systems

In this section the processes for recommendation calculation will be described which are used to predict preference values for the current user. The presented calculation processes are oriented to the differentiation of functional input sources (see section 2.2). The subsequent section 3.3 lists calculation functions which can be utilized within the calculation processes presented here.

Calculation process of content-based systems. Content-based systems calculate preference values based on objective properties, the features. Using these features an item-to-item-correlation is constructed based on the preferences of the current user [5]. The objective is to identify unrated information artifacts which is alike rated information (information the system has preference values for).

The calculation process of content-based systems is composed of two phases, as shown in Fig. 1 (left). In the first phase 'feature extraction' objective properties are extracted for the information artifacts, with the aim of identifying the best describing characteristics of this artifact. In the second phase 'filtering and preference value prediction' a heuristic or model-based mathematical function is used to calculate the similarity of informational artifacts. The applied functions can be distinguished in exact-match and best-match approaches.

Calculation process of collaboration-based systems. Collaboration-based recommendation systems recommend information artifacts to the current user based on community data, the user/usage data of other users. The process is composed of three phases, as shown in Fig. 1 (right). In the first phase a user-to-user-correlation is constructed based on a comparison of acquired preferences of these users (or user account). The aim of the second phase is to identify a homogeneous neighborhood group of mostly similar users. In the last phase the preference value predictions are calculated. The resulting recommendations for the current user are these information artifacts which are preferred by other users which have mostly similar preferences to the current one [5,13,18].

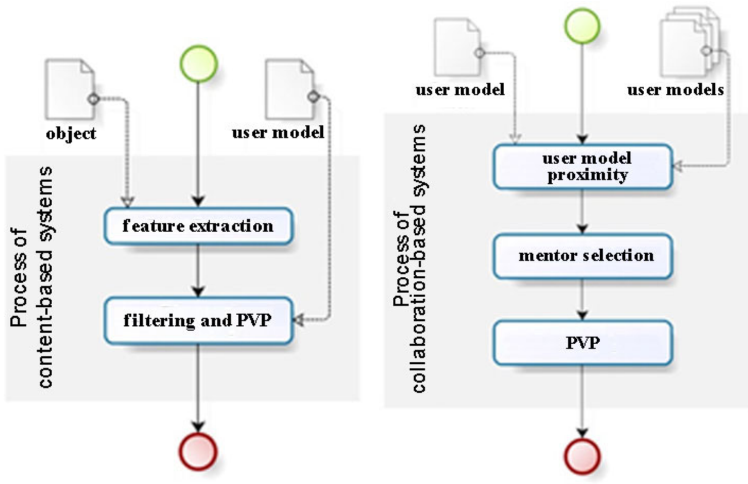


Fig. 1. Calculation processes for content-based (left) and collaboration based (right) systems

Hybrid systems. Hybrid recommendation systems combine the approaches of content-based and collaboration-based processes. The objective of the combination is to reduce weaknesses and problems when applying the above processes separately [3]. [1] identified four categories of combinations for hybrid systems:

(i) Combination of results: both processes are integrated in independently, the resulting preference predictions are combined afterwards.

(ii) Content-based aspects in collaboration-based methods: content-based user models are utilized to calculate homogeneous neighborhood groups.

(iii) Collaboration-based aspects in content-based methods: the predominant approach is a dimension reduction of content-based user models based on collaboratively constructed neighborhood groups.

(iv) Unified model: both processes are modeled in one joint model which is used to calculate the preference value predictions. This model can be probabilistic, rule-based, or knowledge-based.

3.3 Computation Functions in the Recommendation Process

In the phases of the recommendation calculation processes mathematical functions can be applied. In this section these some applicable functions will be introduced, which is necessary to identify the information flows in the comprehensive reference model for recommender systems.

Feature extraction. For the phase of feature extraction the text-based approach TF-IDF is mentioned which identifies relevant terms of documents and their significances. Using latent semantic analysis or respectively latent semantic indexing the quality of this feature extraction can be increased [1].

Proximity calculation. To calculate the proximity between user models five strategies can be distinguished [1,9,12,13,14,16,17]:

1. 1. distance-based: Lq-distance, Mahalanobis-distance,
2. 2. correlation-based: (constrained/weighted) pearson correlation,
3. 3. vector-space-based : cosine-based approach, adjusted cosine similariy, TF-IDF-schema, rocchio algorithm,
4. 4. rule-based: boolean matching, decision trees, association rule discovery, and
5. 5. probabilistic: Hidden Markov Model, dynamic Bayesian networks, Bayesian belief networks, naïve Bayesian classifier, maximum-entropy-method, probabilistic relational model, probabilistic latent semantic analysis.

Classification. A classification leads to groups of 'similar' objects based on a metric function, which is here based on features. Here three approaches can be distinguished [4,6,19,20]:

1. 1. Nearest Neighbor: kNN,
2. 2. linear classifier: (modified) least squares, Support Vector Mashines, modified naïve Bayes, and
3. 3. Cluster-algorithms: Tree-Clustering, Expectation-Maximization-Algorithm, K-Means, Gibbs Sampling, Monte-Carlo-Methods, Horting-Algorithm

Aggregation. Aggregation functions unite multiple values to one. Approaches are the simple average (means), weighed sum, adjusted weighed sum or the regression model. A similar approach are providing the (bi-polar) slope one-algorithms [1,13,14].

4 Comprehensive Reference Model for Recommender Systems

In this section the comprehensive reference model for recommender systems will be presented. The model is based on the Recommendation Calculation Processes for Recommender Systems (see section 3.2) and summarizes the main aspects of related work reference models (see section 3.1). Furthermore the information flows in the reference model will be identified.

4.1 Aggregation of Existing Process Models

The existing process models are described in section 3.2. To achieve the comprehensive reference model, the processes of content- and collaborative-based recommendation systems had been combined and extended by the transitions and phases of hybrid systems.

The resulting comprehensive reference model for recommendation systems is presented in Fig. 2. The entry point for this model is the green node at the upper side, the endpoint with the resulting preference value predictions is the red node at the bottom side. The yellow squares are decision points which path has to be followed. The decision itself is mainly determined by design conditions. Input data for the reference model are the objects itself, or to be precise, features of these objects for which the preference value predictions are to be calculated, as well as user models, containing processed user information (like interactions or selection/buy decisions). These user models are differentiated in the user model for the current user for whom the preference values are to predict, and the user models of other users.

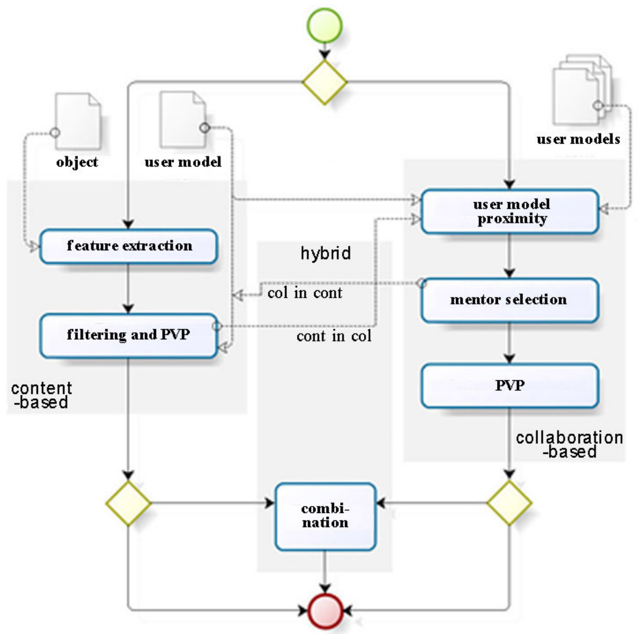


Fig. 2. Comprehensive reference model for recommendation systems

In Fig. 2 the process of content-based recommendation systems is presented in the left grey box, whereas the process of collaboration-based recommendation systems is visible in the right grey box. In between these two boxes transitions for possible hybrid combinations are displayed to illustrate the intersections between these two classes of recommendation systems. The smaller grey box at the bottom side of the model represents the class of hybrid recommendation systems which combines the results the two processes calculated independently. The residual category of unified models (see section 3.2) is not distanced in form of transitions or phases, but is a form auf instantiation of the whole process model.

4.2 Data Flow in the Transitions of the Reference Model

In Fig. 2 some transitions are displayed between the phases of the recommendation processes. These been examined for their data flow. Therefore, the computation functions described in section 3.3, which may be executed in the phases, had been taken into account. These computation functions expect specific input data and process them to a specific output data. Therefore these data is passed between the connected phases when recommendations are calculated. In Fig. 3 these annotated data flow in the transition is presented.

The features of the objects may be extracted using feature extraction algorithms. The resulting feature vector is passed on to the phase ‘filtering and preference value prediction’. In conjunction with the feature vector of the current users’ user model distance values are calculated using proximity calculation algorithms. The results of the content-based recommendation process are the preference value predictions.

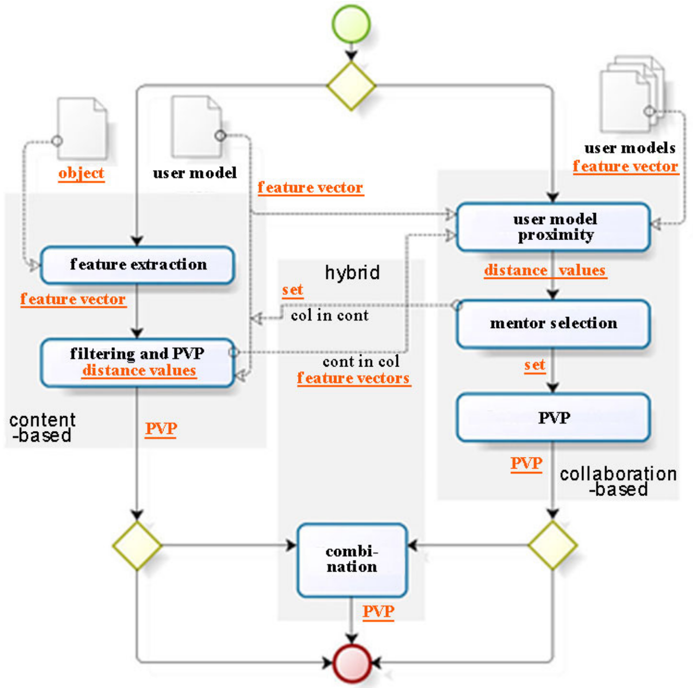


Fig. 3. Comprehensive reference model for recommendation systems extended by annotations of the data flow at the transitions

For the collaboration-based recommendation process both feature vectors of the current users' user model and the user models of other users are the input for the phase 'user model proximity' calculation. Here again proximity calculation algorithms are applied which generate distance values. These values flow into the phase 'mentor selection' where classification algorithms are applied. Using the resulting user model set describing users likewise similar preferences, in the phase 'preference value prediction' these predictions are calculated.

For hybrid systems combining the processes the one approach is to hand over the set of mentors to the content-specific phase 'filtering and preference value prediction'. The other approach of bringing aspect of one process into the other process is to pass the resulting feature vector of the current users' user model (containing the process result of the objects feature vector) to the collaboration-specific phase 'user model proximity calculation'. The category of combining the results uses aggregation functions to merge the calculated preference value predictions.

5 Conclusion

For the success of recommendation systems in applications it is crucial the current user accepts recommended objects. The acceptance depends on the traceability, which in turn depends on the information the user may perceive. Using the model presented in this paper those information of the recommendation calculation process can be identified which should be depicted to the user to enhance traceability. So the user can better understand why objects had been recommended. This may increase the success and returns for the vendors utilizing the system.

In this paper we presented a comprehensive reference model for recommender systems. This reference model conjuncts the processes of content-based and collaboration-based recommendation systems and extends these processes with the transitions and phases of the category of hybrid systems. Furthermore, the algorithms which can be applied in the phases of the model had been examined to identify the data flow between these phases.

The resulting reference model presents the state of the art approaches of recommender systems to calculate preference value predictions. Using this reference model, developers can systematically construct new recommendation systems or develop further improvements and extensions of existing recommendation systems.

Acknowledgments. The here described work was developed as a part of the Core-Technology-Cluster for Innovative User Interfaces and Visualizations of the THESEUS Program, a 60-month program partially funded by the German Federal Ministry of Economics and Technology.

References

1. Adomavicius, G., Alexander, T.: Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions. IEEE Trans. on Knowledge and Data Engineering 17(6), 734–749 (2005)
2. Anick, P.: Using terminological feedback for web search refinement: a log-based study. In: SIGIR 2003: Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 88–95. ACM Press, New York (2003)
3. Balabanovic, M., Yoav, S.: Fab: content-based, collaborative recommendation. Communications of the ACM 40(3), 66–72 (1997)
4. Breese, J.S., Heckerman, D., Kadie, C.: Empirical Analysis of Predictive Algorithms for Collaborative Filtering. In: Proceedings of the 14th Conference of Uncertainty in Artificial Intelligence, pp. 43–52. Morgan Kaufmann, San Francisco (1998)
5. Burke, R.: Hybrid Recommender Systems: Survey and Experiments. User Modeling and User-Adapted Interaction 12(4), 331–370 (2002)
6. Iyengar, V.S., Zhang, T.: Empirical Study of Recommender Systems Using Linear Classifiers. In: Cheung, D., Williams, G.J., Li, Q. (eds.) PAKDD 2001. LNCS (LNAI), vol. 2035, pp. 16–27. Springer, Heidelberg (2001)
7. McCarthy, K., Salamo, M., Coyle, L., Mc-Ginty, L., Smyth, B., Nixon, P.: Group recommender systems: a critiquing based approach. In: Paris, C., Sidner, C.L. (eds.) IUI, pp. 267–269. ACM Press, New York (2006)
8. Micarelli, A., Gasparetti, F., Sciarrone, F., Gauch, S.: Personalized Search on the World Wide Web. In: Brusilovsky, P., Kobsa, A., Nejdl, W. (eds.) Adaptive Web 2007. LNCS, vol. 4321, pp. 195–230. Springer, Heidelberg (2007)
9. Min, Sunghwan und Ingoo Han: Optimizing Collaborative Filtering Recommender Systems. In: Szczepaniak, Piotr S., Janusz Kacprzyk, Adam Niewiadomski (Ed.): AWIC, LNCS Vol. 3528, pp. 313-319, Springer, (2005).

10. Neumann, A.W.: Recommender Systems for Scientific and Technical Information Providers. Doctor thesis, University of Karlsruhe, Germany (2008)
11. Resnick, P., Varian, H.R.: Recommender Systems – Introduction to the Special Section. Communications of the ACM 40(3), 56–58 (1997)
12. Rocchio, J.J.: Relevance Feedback in Information Retrieval. In: Salton, G. (ed.) The SMART Retrieval System – Experiments in Automatic Document Processing. Prentice Hall, Englewood (1971)
13. Runte, M.: Personalisierung im Internet - Individualisierte Angebote mit Collaborative Filtering. Doctor thesis, University of Kiel, Germany (2000)
14. Sarwar, B., Karypis, G., Konstan, J., Riedl, J.: Item-based Collaborative Filtering Recommendation Algorithms. In: Proceedings of the 10th International World Wide Web Conference, pp. 285–295 (2001)
15. Schafer, J.B., Konstan, J.A., Riedl, J.: E-Commerce Recommendation Applications. Data Mining and Knowledge Discovery 5(1-2), 115–153 (2001)
16. Shardanand, U., Maes, P.: Social Information Filtering: Algorithms for Automating 'Word of Mouth'. In: Proceedings of the Conference on Human Factors in Computing Systems (1995)
17. Su, X., Khoshgoftaar, T.M.: A Survey of Collaborative Filtering Techniques. Advances in Artificial Intelligence, 1–20 (2009)
18. Terveen, L., Hill, W.: Beyond Recommender Systems: Helping People Help Each Other. In: Carroll, J. (ed.) HCI in the New Millenium. Addison-Wesley, Reading (2001)
19. Ungar, L., Foster, D.: Clustering Methods for Collaborative Filtering. In: Workshop on Recommendation Systems. AAAI Press, Menlo Park (1998)
20. Zhang, T., Iyengar, V.S.: Recommender systems using linear classifiers. J. Mach. Learn. Res. 2, 313–334 (2002)

Dynamic Interface Reconfiguration Based on Different Ontological Relations*

Elisa Chiabrando, Roberto Furnari, Pierluigi Grillo,
Silvia Likavec, and Ilaria Lombardi

Università di Torino, Dipartimento di Informatica, Torino, Italy
{chiabrando,grillo,furnari,likavec,lombardi}@di.unito.it

Abstract. This paper presents a framework and a prototype implementation of a system that integrates ontological knowledge base with the services for the users in mobility. Our system tries to help users navigate the domain of gastronomy, offering a possibility to explore the ontological base consisting of many concepts and relations. We present the logic and technical realization of the modules responsible for knowledge base interrogation and for presenting the obtained information to a user. We also provide a brief description of the implemented interface and user evaluation.

1 Introduction

The Internet started as a plain static information source, but, nowadays, it is rapidly changing to a dynamic interactive platform where users actively produce content, create social networks and interact with each other. It became possible to obtain information on the move, in the moment when it is needed, due to increased number of mobile devices on the market and the availability of Internet access in virtually any situation.

Huge amount of information available on the Web is organized in different modes, with ontologies being used very often to model various domains, since they provide structured information on which reasoning is readily available.

Visualizing vast amounts of heterogeneous information in social networks is a challenging task. Especially in the "Mixed social networks", where objects of the domain and users of the network can interact with each other. In our framework, we create exactly this kind of network, enabling the interaction among people and physical objects as indistinguishable parts of the society, as well as the interaction of people with virtual avatars of the objects that continue to exist even when the user is not in direct contact with the physical objects.

In particular, the framework described in this paper is a part of the PIEMONTE Project [9], in which a Social Networking Application in the domain of gastronomy is being developed by integrating social networking and augmented reality tools. In this particular domain, there is a growing need for information

* This work has been supported by PIEMONTE Project - People Interaction with Enhanced Multimodal Objects for a New Territory Experience.

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 538–547, 2011.

about certain products, local producers, production technologies, as well as about similar users and their opinions. Our application developed for Apple iPhone tries to support producers and end-users, meet the expectations of users in mobility and enhance their mutual interaction and overall experience.

We found sophisticated solutions to present the users with different aspects of the domain in a gradual way, enabling them to profit from the great amount of information available. To this aim we developed a new interaction model called the "wheel". The main contributions of our framework are the following:

- a strategy to select the content to be presented to the user (content of the "wheel");
- a novel visualization technique for presenting the relevant information (appearance of the "wheel");
- a novel interaction technique for multi-touch mobile devices (spinning the "wheel").

In this paper we concentrate on the technical realization of the framework and the wheel content selection strategy based on reasoning over domain knowledge base and analyzing the user model. We briefly describe the visualization technique for presenting the information to the user, the details of which can be found in [1].

The paper is organized as follows. In Section 2 we touch upon some related work in the field. We describe the architecture we developed for our framework in Section 3, followed by the ontological organization of the domain given in Section 4. In Section 5 we discuss how the relations among the objects of the system and between users and the objects of the system are established. A brief overview of the developed interface is given in Section 6. In Section 7 we summarize the evaluation results and conclude in Section 8.

2 Related Work

The objective of SEWASIE project[1] is to provide users with uniform access to heterogeneous ontological data. In [3] authors describe the basics of an intelligent user interface used to query such data. Our aim is different, we want to provide users with an opportunity to navigate through a vast space of information in a ludic and intuitive way.

An example of using mobile phones to access ontologically organized data is given in SmartWeb project.[2] In [13] a context-aware, multi-modal mobile application for the Semantic Web is developed. The users interact with the application using smartphones, asking questions using speech and gesture. In [12] the users can obtain answers to their questions in a specific domain as multimodal presentations based on graph visualization and ontological navigation.

Another framework for ontology navigation is presented in [5]. Users can freely explore the ontology or query the information system, neither needing to have the

[1] `http://www.sewasie.org/`

[2] `http://www.smartweb-project.de/`

knowledge about the internal organization of the data, nor about the vocabulary used to describe the domain. Their software called Query Tool consists of three components: the query logic used to reason over the ontology, the engine for natural language generation, and user interface.

With respect to the mentioned works, although in our framework ontologies play a central role in the domain representation, the ontological knowledge is enriched by knowledge from other sources and offered to the user by means of a Resource Oriented Architecture.

The work in [4] presents a hardware kit (electronic bracelet with sensors, RFID reader and accelerometer), which allows the user to interact with real objects. Using hand gestures the user can browse vocal menu provided by objects and select the actions to perform.

[11] describes a Physical Mobile Interaction Framework (PMIF) for the communication between a device (e.g. smartphone) and objects in the environment. These objects can communicate with the device "passively" using RFID, NFC, visual markers or bluetooth connection, or they can communicate "actively" having their part of framework-client (PMIF) installed.

On the other hand, in our framework communication with objects is much simpler, since we prefer to use a common device, such as a smartphone, to achieve interaction with objects where there is no need to modify them by inserting codes or installing sensors.

With respect to the software architecture we used, the benefits of using a Resource Oriented Architecture are highlighted in [6] in a Web of Things context; although in a different context, similar consideration can be done in our work, where resources represent a sort of avatars of real things that tell the users about their properties and their relations.

3 System Design and Implementation

In the following section we briefly describe the architecture we developed for our framework. Figure 1 depicts the layered architecture of our framework, where client applications ask for particular aspects of domain knowledge and the service infrastructure provides the information required, enabling the clients to browse the domain knowledge gradually and smoothly.

The requested information is provided to the client applications by means of a set of related resources organized in a Resource Oriented Architecture [10]. In such a perspective the items composing the domain knowledge can be easily inquired about their own features and relations with the other domain resources. Following these links between resources it is possible to browse the network of domain items and discover new parts of the domain in a consistent way.

The Resource Oriented Architecture in our framework is managed by means of the *Service Logic Layer*, the core component of the server side infrastructure, while the *Exposure Layer* plays the role of a controller that manages the communication between the *Service Logic Layer* and the client applications from the *Application Layer*, providing also a monitoring service.

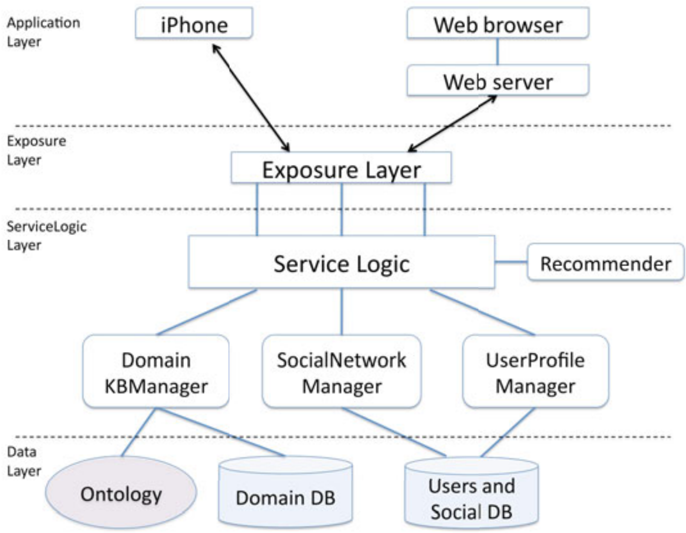


Fig. 1. System architecture

The *Service Logic Layer* organizes the different knowledge sources present in the *Data Layer* synthesizing a uniform representation of the resources. When requested, specialized modules (*UserProfile Manager* and *Recommender*) are exploited to customize the service depending on the user interests.

In the *Data Layer*, two main kinds of knowledge exist: a domain knowledge where the domain items and their relations are represented (see Section 4 for more details), and a social knowledge where the information about the users and their social actions is stored. In particular, the *DomainKB Manager* offers an integrated view of the two data sources used in the *Data Layer* in order to represent the domain knowledge: the ontology and a database storing a sort of master data set[3] about domain items.

The architecture has been realized using Java Environment where the Restlet tool[4] and Apache HTTP libraries collaborate in order to answer the HTTP requests from clients. The data are stored using an OWL ontology integrated with a relational MySQL database. The *DomainKB Manager* uses the D2RQ tool[5] [2] to provide a uniform virtual data source that can be directly inquired by means of SPARQL query. Mechanisms of caching are implemented by the *Service Logic* in order to optimize the request answering process. The information toward a client is serialized in JSON streams, which makes the interoperation

[3] By master data set we mean long text description, contact information, pictures and video references, etc. This is all the information about the domain items that does not benefit from the ontological representation since it is excluded from the inference reasoning.

[4] `http://www.restlet.org/`

[5] D2RQ allows to view a part of a relational database as a set of OWL individuals.

with different client types smooth and seamless. In fact, although the focus of this paper is on the Apple iPhone mobile client, different clients like Web Servers are also served by this architecture. Our application for Apple iPhone has its complement in a form of a web application that provides the producers with the opportunity to enter the gastronomical world and make their products known.

4 Ontological Organization of the Domain

Our framework uses ontologies to represent the domain knowledge as a machine recognizable content. Adopting ontologies as knowledge bases enables smooth description and integration of data, as well as queries and reasoning on them. Ontologies describe the domain items by expressly representing their properties and the relations among them, highlighting their similarities and differences. Ontology tools, such as subclasses and restrictions, provide a well structured and organized description of the domain where new data can be seamlessly added. Moreover, ontology reasoning makes available, by means of inference, properties and information that are implicit in the knowledge base. In this project we use OWL 2[6] - the Web Ontology Language, that extends the previous OWL definition with support for datatypes with ranges and metamodeling.

The domain of our application is gastronomy where we deal with different kinds of domain items such as products, actors related to products (producers, farms, restaurants), as well as recipes and geographical data. The entities of the domain are modeled as elements of the ontology containing very general classes such as `Wine` or `Cheese`, as well as very specific ones, such as `Dolcetto d'Asti di Vittorio Ubertone` or `Toma di Lanzo di Marino Dardino`.

In order to cover this wide range of knowledge, the model we use is composed of several ontologies, specifying different domain parts, that are imported by the global upper ontology. In particular, we use the following ontologies:

– **Taste Ontology** is the global upper ontology which imports the other ontologies and where the connections between objects belonging to different ontologies are made explicit.
– **Products Ontology** is the most populated ontology which describes the products of the domain. It includes other sub-ontologies that detail particular product branches. For example, the Wine Ontology and Cheese Ontology classify and describe wines and cheeses in great detail. Items are grouped by kinds and properties related to their organoleptic features, production procedures and membership in particular quality consortiums.

  Furthermore, a set of SWRL[7] rules are defined on this ontology to relate foods based on their properties: in particular, associations between wines and cheeses are defined.

[6] `http://www.w3.org/TR/owl2-overview/`
[7] `http://www.w3.org/Submission/SWRL/`

- **Actors Ontology** stores the actors playing active roles in our domain. Both virtual actors, such as food producing companies, farms, restaurants or shops, as well as physical actors, such as producers, restaurant chefs and sellers are present. The actors are related to products by properties: producers produce products, cooks and restaurants use them etc.
- **Geo Ontology** is a taxonomy of the geographical places of the domain (places of origin of products, places where the products are sold etc.). It includes the administrative regions, as well as the areas that do not correspond to administrative divisions but are widely used to identify particular places. For instance valleys like `Valsusa` or `Val di Lanzo` are the names referring to the regions included in the district of Turin and grouping several municipalities.
- **Recipes Ontology** with the ingredients for each recipe, together with the cooking instructions. Note that the ingredients of a recipe can be recipes itself (e.g sauces or creams), which makes the relationship among ontology elements particularly interesting.

As mentioned in Section 3 the domain knowledge, provided by ontologies, is enriched with the information from the data base. This results in a complete domain description where further investigations about relations among domain items take place.

5 Inferring Object Relations

Our framework, like other Social Web applications, encourages users to provide content, such as new entities of the domain, tags, ratings and comments. Since this causes a considerable growth of the available information which risks to overwhelm the user, a strategy to limit and tailor the information presented to the user is of major importance.

Our *Recommender module* is in charge of adapting the content presented to the user by filtering and, most notably, ranking the entities, taking into account user preferences and the relations among entities (objects and people) of the system. Such relations, which are an essential part of the mixed social network, are described in the ontology or can arise from users' actions.

The most natural way to define a link between two instances in the ontology is by using the *ObjectProperty*. For example, the class `Product` can have an ObjectProperty *producedBy* whose range is the class `Producer`. Moreover, relations between entities and places, or between products and actors in the production chain, as well as restaurants, shops, recipes and chefs exist. We also define SWRL rules to express more sophisticated associations based on the property values of the instances. For example, a set of rules pairs cheeses and wines to be served together, considering their organoleptic properties.

An interesting aspect of the system is that it can infer relations based on user actions. Apart from explicit associations, for example, if a user chooses another person as a friend or if she adds a product as a favorite, the system can recognize implicit connections analyzing tags and comments. Hence two entities can be linked if one is tagged with the name of the other one, or if they are tagged with the same label by a significant number of users.

This kind of analysis allows for dynamic inferring of user preferences and of users' similarity. On the one hand, the actions the user performs on entities are claims of interest on them, with an increasing degree from rating, to tagging and commenting. On the other hand, the system can infer a relationship between two users who perform similar actions on the same objects as this reveals a similarity between them.

6 Content Presentation

As stated above, the interaction with the user is carried out using a mobile device (Apple iPhone) where the user can explore the domain content by using the *WantEat* application. *WantEat* is an application which unifies the default elements of an iPhone application (action sheet, navigation bar, etc.) with a new navigation model, called the *wheel*.

The wheel consists of the central object, which is the *focus of the wheel*, and four different *Sectors* around it (See Figure 2(a)). Each Sector groups objects related with the central one by particular relations. These objects are the result of a *Service Logic* query (see Section 3) and the relations are those inferred by the system (as described in Section 5). For instance, a Sector labeled *Territory* will collect all the objects that are connected to the focus of the wheel with a relation having to do with the territory; e.g. places where a product grows, shops where a product is sold, production companies where it is produced. Due to the heterogeneity of the domain items we deal with, different *sector configurations*[8] are possible. In particular, the sector configuration used depends on the kind of the object in the focus, and changes dynamically as the central object changes.

Using the wheel a user can browse and discover the items of the domain in a gradual and non-confusing way. Starting from the wheel of an object, the user can select a Sector, expand it and view the objects belonging to it. A Sector can contain a large number of items, due to the different possible relations between the central element and other domain elements. The user can further explore the different Sector elements, and the different relations these objects have with the object in the center by using an innovative interaction technique called "spinning the wheel", i.e. scrolling or rotating the elements in the Sector. If the user is interested in a particular item, this can become the new focus by dragging it to the center of the wheel, and the wheel configuration changes accordingly.

For example, Figure 2(a) depicts the *wheel* for the product `Toma di Lanzo` (a cheese), where the following sector configuration is used: *territory* (e.g. places where it is produced), *people* (e.g. users that tagged or voted it), *cuisine* (e.g. recipes using it) and *products* (e.g. similar products). Figure 2(b) shows the wheel where the Sector Territory is expanded. When the user changes the focus by dragging the shop `El Canton dij Formagg` from its Sector to the center

[8] We call *sector configuration* the name, the position and the features of the four Sectors forming the wheel.

(a) (b) (c) (d)

Fig. 2. Dynamic change of the sectors

Fig. 3. Dynamic interface in video domain: iPad application

of the wheel (see Figure 2(b)), the sector configuration changes automatically: *cuisine* and *products* are replaced with *products* and *shops* (see Figure 2(c)).

If the user is not interested in a particular type of a relation or a type of an object, the interface provides different criteria to filter the items in the sectors (the buttons in the lower portion of the screen). In particular, it is possible to show only the items related to the *Slow Food* organization[9] [8], to show the items local with respect to the user position or to hide all the elements of a kind (products, producers, restaurants, etc.).

All the objects of the interface can be *touched* to get more details. It is also possible to investigate a particular relation in order to find out why two objects are related, i.e. which relation connects them (see Figure 2(d)).

[9] `http://www.slowfood.com`

In addition to interaction with the wheel, users can perform various Web 2.0 actions: tagging, voting, commenting or putting into favorites the items they find interesting.

Our novel interface can be used in other multimedia contexts. We extended the *WantEat* iPhone application with the *WantEatVideo* version for Apple iPad in the video domain (see Figure 3). In this case, the videos (of the production process of certain products, for example) are connected with the objects of the wheel and change dynamically as soon as the objects in the sectors change.

7 Evaluation

The first proposal of our framework was tested at the early stages in order to incorporate user feedback into the final prototype. After testing the robustness of the system and the consistency of the inserted data, we performed a large scale user evaluation in October 2010 during the International Food Fair "Salone del Gusto" in Turin, Italy. More than 600 users were divided into four groups depending on their age and technology usage. They used Apple iPhones to explore the objects of our domain. We wanted to test the intuitiveness of our approach, easiness of usage and usefulness of the information obtained.

More than half of the tested users found the application easy and enjoyable to use and the information obtained useful. In addition, valuable suggestions were gathered about possible improvements and additional services (product prices availability, possibility to book a hotel/restaurant or buy a product etc.)

8 Conclusions

We described a novel framework for representing vast amounts of heterogeneous domain knowledge with innovative user interface and interaction technique.

The interaction experience that the application provides and the information that can be discovered is the important aspect to keep in mind. The information is presented incrementally, providing the user with non-confusing results.

The user interface was particularly designed for users in mobility, using Apple iPhone as interaction device, where we tried to introduce some novel interaction elements, while preserving many typical aspects of iPhone applications. Hence, the user has the opportunity to explore new interaction techniques and comprehensive domain knowledge, while at the same time being relaxed in the familiar environment. These new interaction metaphors have the potential of being employed in other similar applications.

Our framework can be applied to broad spectrum of different domains, reusing the approach to knowledge representation and organization, as well as a way of inferring relations between objects and people, and people themselves. One immediate future research direction that comes to mind is using a sort of similarity measure (see [7]) to asses how similar the objects in the ontology are. Creating a so called "Mixed social network" where real and virtual counterparts of objects and people can interact at any time is a powerful concept, influence of which should not be underestimated. It is a concept that offers many directions for future enhancements and enlargements.

References

1. Biamino, G., Grillo, P., Lombardi, I., Marcengo, A., Rapp, A., Simeoni, R., Vernero, F.: "The Wheel": an innovative visual model for interacting with a social web of things. In: Workshop on Visual Interfaces to the Social and Semantic Web, VISSW 2011 (part of IUI 2011), Palo Alto, CA, USA (2011) (accepted for publication)
2. Bizer, C., Seaborne, A.: D2RQ - Treating Non-RDF Databases as Virtual RDF Graphs. In: 3rd International Semantic Web Conference, ISWC 2004 (Poster), http://www4.wiwiss.fu-berlin.de/bizer/d2rq/
3. Catarci, T., Dongilli, P., Mascio, T.D., Franconi, E., Santucci, G., Tessaris, S.: An ontology based visual tool for query formulation support. In: Chung, S. (ed.) OTM-WS 2003. LNCS, vol. 2889, pp. 32–33. Springer, Heidelberg (2003)
4. Feldman, A., Tapia, E.M., Sadi, S., Maes, P., Schmandt, C.: ReachMedia: On-the-move interaction with everyday objects. In: 9th IEEE International Symposium on Wearable Computers, ISWC 2005, pp. 52–59 (2005)
5. Franconi, E., Guagliardo, P., Trevisan, M.: An intelligent query interface based on ontology navigation. In: Workshop on Visual Interfaces to the Social and Semantic Web, VISSW 2010 (2010)
6. Guinard, D., Trifa, V., Mattern, F., Wilde, E.: From the Internet of Things to the Web of Things: Resource Oriented Architecture and Best Practices. In: Architecting the Internet of Things. Springer, Heidelberg (2011)
7. Likavec, S., Lombardi, I., Nantiat, A., Picardi, C., Theseider Dupré, D.: Threading facts into a collective narrative world. In: Aylett, R., Lim, M.Y., Louchart, S., Petta, P., Riedl, M. (eds.) ICIDS 2010. LNCS, vol. 6432, pp. 86–97. Springer, Heidelberg (2010)
8. Petrini, C.: Slow Food Nation: Why Our Food Should Be Good, Clean, and Fair. Rizzoli (2007)
9. PIEMONTE Team. WantEat: interacting with social networks of intelligent things and people in the world of enogastronomy. In: Interacting with Smart Objects Workshop 2011 (part of IUI 2011), Palo Alto, CA, USA (2011) (accepted for publication)
10. Richardson, L., Ruby, S.: RESTful web services. In: Media, O, ed. (2007), http://www4.wiwiss.fu-berlin.de/bizer/d2rq/
11. Rukzio, E., Wetzstein, S., Schmidt, A.: A Framework for Mobile Interactions with the Physical World. In: Invited Paper Special Session "Simplification of User Access to Ubiquitous ICT Services" in Wireless Personal Multimedia Communication, WPMC 2005 (2005)
12. Sonntag, D., Heim, P.: Semantic graph visualisation for mobile semantic web interfaces. In: Hertzberg, J., Beetz, M., Englert, R. (eds.) KI 2007. LNCS (LNAI), vol. 4667, pp. 506–509. Springer, Heidelberg (2007)
13. Wahlster, W.: SmartWeb: Mobile Applications of the Semantic Web. In: GI Jahrestagung (1). LNI, vol. 50, pp. 26–27. GI (2004)

Analysis of Content Filtering Software*

Woongryul Jeon[1], Youngsook Lee[2], and Dongho Won[1]

[1] School of Information and Communication Engineering, Sungkyunkwan University, Korea
{Wrjeon,dhwon}@security.re.kr
[2] Department of Cyber Investigation Police, Howon University, Korea
ysooklee@howon.ac.kr

Abstract. The openness of web allows any user to easily get information anytime and anywhere. However, as well as the advantage of easy access to useful information, the web has disadvantages of providing users with harmful information indiscriminately. Some information, such as adult content, is not appropriate for children. To protect children from adult contents, many types of filtering software have been distributed. However, filtering software cannot entirely block adult contents, because it has some functional limitations. In this paper, we analyze several filtering software and discuss their limitations. Furthermore, we propose the essential functionalities for filtering software.

Keywords: Evaluation, Filtering Software, Adult Contents Filtering.

1 Introduction

Openness of web allows any user to easily get information anytime and anywhere. However, as well as advantages of easy access to useful information, the web has disadvantages of providing users with harmful information indiscriminately. For example, some information, such as adult contents, is not appropriate for children. In fact, there is clear evidence that many children who use the internet are exposed to pornographic websites. According to the Korea Communications Standards Commissions, 83.4% of respondents encounter adult contents on the web[1, 2]. In case of children, they cannot determine what is right and wrong, because they are too young. For this reason, exposure of adult contents to children can cause crimes. Actually, many sexual crimes have caused by children around the world.

To block adult contents from children, several filtering software is distributed. The filtering software is one of the effective methods to block adult contents from children. General functionalities of the filtering software are as follows: First, the filtering software blocks access to adult contents according to predefined criteria. This constitutes the basic functionality of filtering software. Second, filtering software allows parents to make limitation for their computer about how much time or at what time children can use the computer or internet. Third, the filtering software allows parents to audit their computer, so they can check the online/offline activity of their children.

* This research was supported by the MKE(The Ministry of Knowledge Economy), Korea, under the "ITRC" support program supervised by the NIPA(National IT Industry Promotion Agency)" (NIPA-2011-C1090-1001-0004).

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 548–557, 2011.

The audit functionality of the filtering software can provide records of all auditable events, such as visited websites, executed programs including screen shots of programs, duration of computer usage, etc.

However, filtering software cannot be a perfect solution for blocking adult contents. In fact, some filtering functionalities can be bypassed because of its implementation and, some filtering software only operates in a specific environment.

Thus, in this paper, we analyze several filtering software, and its effectiveness. Furthermore, we discuss the limitations of filtering software and propose essential requirements of the filtering software.

The remainder of this paper as follows. In section 2, we define adult contents and in section 3, we provide criteria of evaluation. The result of evaluation is given in section 4, and in section 5, we propose requirements and essential functionalities of filtering software based on evaluation. Finally, conclusion is given in section 6.

2 Definition of Adult Contents

Before evaluation of filtering software, we have to define what adult contents are. To define this, we refer to a standard which is defined by the ICRA (Internet Content Rating Association). The ICRA is an international non-profit organization of United States of America. The purpose of the ICRA is to help users find desired and trusted content and to filter out undesired content for themselves or their children.

The ICRA does not impose content ratings, but allows web masters and digital content creators to self-label their content into categories such as nudity, sex, language, violence, and other potentially harmful material and chat. Because the content rating relies on the contents developer, it is possible that part of the content or site may depict nudity, but it is not rated. Thus, contents consumers, parents or children, have to make decisions about which contents they find appropriate. The ICRA classes harmful information into several categories, Nudity & Sexual Material, Chat, Languages and etc[4]. In ICRA, young children means under the age of 12.

Table 1. Adult Contents

Category	*Content Level*
Chat	Any immoderate chat must be blocked.
Language	Any explicit crude words must be blocked.
Nudity & Sexual Material	Any genitals, breasts, buttocks, and sexual acts must be blocked except artistic, educational, medical, and sports context.
Violence	Any blood and killing of human beings or animals must be blocked.
Other Topics	Any context that promotes drug, tobacco, alcohol, weapon usage and violence must be blocked.

The Nudity & Sexual Material and Sports categories have a context field. Context fields are necessary to allow intended contents such as articles, educational, medical and sports contents. For example, Botticelli’s “Birth of Venus” contains a female breast and so it may seem that this content is not appropriate for children by the ICRA criterion. However, this content is regarded as an artistic painting so it is suitable for young children. For this reason, context fields are included in the Nudity & Sexual Material and Sports categories.

In this paper, we purpose to block adult contents from children who are under the age of 12 according to the ICRA criteria. Table 1 shows definition of adult contents.

3 Base of Evaluation

3.1 Selection of Filtering Software

In this section, before analysis and testing of filtering software, we select sample filtering software. There are various filtering software and their functionalities are also various. To insure reliability, we have to select filtering software carefully. Thus we adopt a number of criteria to select filtering software. Table 2 describes the criteria for selecting software.

Table 2. Criteria for Filtering Software

Criteria	*Description*
Various Platforms	Filtering software should support various common platforms e.g. Windows 98/XP/Vista, Linux, Mac OS etc.
Various Browsers	Filtering software should support various common web browsers, e.g. Internet Explorer, Mozilla FireFox, Google Chrome etc.
Various Protocols	If filtering software supports various protocols, then it can be regarded as good, because it can block adult contents from reaching various paths.
Update	Filtering software should have a history of commercial success, so many people should have used it for a long time and it should have been updated recently.

First of all, to obtain objective results, the filtering software has to be common and used widely for a long time. Support for various browsers and platforms can be a good criterion for determining whether or not it is used widely. Support for various protocols is an optional criterion for determining whether or not filtering software is good enough, because children can get adult contents such as messenger, email, voice chatting etc. According to this rule, we select five software, and their descriptions are given in section 4.

3.2 Sample Size

The main purpose of filtering software is to block adult websites from children. However, because there are millions of adult websites around the world, it is hard for filtering software to entirely block all harmful websites. Actually, filtering performance differs according to the filtering software.

Before testing, to insure objective test results, we adopted statistical website sampling. With statistics, we can calculate how many samples are needed with fixed reliability and sampling error. In this paper, we assume 90% reliability and 5% sampling error. In this case, we can calculate the exact sample size, provided we can estimate the total size of the statistical population, i.e., the total number of adult websites. According to the Hankyoreh newspaper in Korea, about 700 thousand websites around the world were opened on the web in 2000. Because this result is ten years old, the present estimate is that there are about one million websites opened on the web.

However, there is not much difference between 700 thousand and one million in terms of calculation of the sample size.

With this total number of adult websites, the appropriate sample size is 271. Thus, we test 400 adult websites and 285 video files for each type of filtering software. Our result has 90% reliability and 5% sampling error.

4 Evaluation of Filtering Software

Following table 3 shows five filtering software tested in this paper.

Table 3. Filtering Software

No	*Software*	*Manufacturer*
1	Parental Control	ICRA
2	Magic Desktop	EasyBits
3	Web Filter PC	Optenet
4	X-Keeper	Jiransoft
5	Momi	Jness

As mentioned in section 2, the ICRA is a nonprofit organization that provides standards for internet content rating. In addition, the ICRA implements a software called parental control that provides filtering functionality. The ICRA is a famous organization and its criteria have been adopted internationally, thus, we selected the parental control program of the ICRA for testing. Web Filter PC is an open source program. X-Keeper and Momi are famous filtering software in the Korea Republic and in particular, these programs can block harmful video files too.

Table 4. Functionality of Filtering Software

Software	*Manufacturer*	*Filtering Functionalities*		
		Website	*Video*	*Finding Files*
Parental Control	ICRA	O	X	X
Magic Desktop	EasyBits	O	X	X
Web Filter PC	Optenet	O	X	X
X-Keeper	Jiransoft	O	O	O
Momi	Jness	O	O	O

The main functionality of filtering software is blocking access to adult contents from children, and this has three types of implementation; filtering websites, blocking video files and finding adult movies, texts and image files. Table 4 shows the functionalities of the seven types of filtering software. In table 4, “O” means that filtering software provides functionality and “X” means vice versa.

4.1 Evaluation of Filtering Software

Test of Filtering Functionalities. The evaluation criteria consisted of three parts, filtering websites, filtering adult movies and finding files. These three factors are the main functionalities that block and filter adult contents from children. Except

X-Keeper and Momi, others do not provide filtering adult movies and finding files, thus filtering adult movies is only tested in X-Keeper and Momi. Table 6 describes the test results of the five filtering software. The test environment is described in Table 5.

Table 5. Test Environment

Operating System	Windows Vista Enterprise K Service Pack 1
Web Browser	Microsoft Explorer 7
Video Player	Gom Player v2.1.16.4631
CPU	Intel Core 2 Duo E5200 2.40 GHz
RAM	2GB
HDD	640GB
Test Date	2009.04.21.

Table 6. Test Result

Software	*Manufacturer*	*Filtering Functionalities*		
		Website	*Video*	*Finding Files*
Parental Control	ICRA	78.22%	None	None
Magic Desktop	EasyBits	100.00%	None	None
Web Filter PC	Optenet	33.68%	None	None
X-Keeper	Jiransoft	53.50%	44.09%	87.90%
Momi	Jness	84.21%	44.09%	36.70%

The test results are dependent on the test date, because the database of each type of software is updated daily, and the available list of websites differs on a daily basis.

Table 7. Network Speed

Filtering Software	*Upload(MB/s)*	*Download(MB/s)*
No Filtering software	**11.29**	**11.67**
Parental Control	11.71	11.71
Magic Desktop	11.58	11.58
Web Filter PC	11.32	11.71
X-Keeper	11.42	11.73
Momi	9.47	11.44

Network Speed. We compare internet speed before and after installation of filtering software. If filtering software affects internet speed, then users may avoid using it. To test internet speed, we used a program called Netmonitoring. This tool provides the average internet speed in 10 seconds. The network test environment uses 100GB LAN. Table 7 shows result of network speed testing. In fact, no filtering software affects network speed.

Resources. If filtering software consumes many resources, it can be a disadvantage of filtering software. Thus, in this section, we examine the resource share of each type of filtering software. In case of the windows operating system, task management provides the resource share for each process and thus we can estimate the resource share of each type of software. The resource share of filtering software is given in table 8.

Table 8. Resource Share of Filtering Software

Filtering Software	*Resource Share(KB)*
Parental Control	1,332
Magic Desktop	2,352
Web Filter PC	4,212
X-Keeper	37,736
Momi	624

The resource share depends on the current situation of the computer, thus the test results can be different. As described in Table 8, X-Keeper shares the most resources among the types of filtering software.

4.2 Limitations of Filtering Software

Filtering based on Database. The main functionality of filtering software is filtering websites, which is based on a database consisting of a set of URLs. To make a database, the manufacturer of the filtering software has to collect the URLs periodically. Using this database, the filtering software can block access to harmful websites in order to compare URLs. If an object URL is already registered in the database, the filtering software blocks immediate access, and generally makes an audit record.

Using a database to block harmful website has two limitations. First, the filtering software cannot block all harmful websites. It is impossible to update all harmful URLs to database, because too many websites, including community sites such as blogs, cafés, and personal homepages, are opened daily on the web. Furthermore, it can be bypassed by renaming URLs. For example, http://sora.net is a harmful website which contains pornography, and this website is already blocked by most filtering software, but http://tosora.info is not blocked, even if this URL indicates the same website, http://sora.net. Second, personal homepages including blogs, cafés, etc., are hard to block with filtering software. In the era of Web 1.0, few people could produce and process information, thus to obtain information, the user had to access portal website, but in the era of Web 2.0, with the advance of IT technology and the spread of internet infrastructure, anybody can produce and process information easily, thus social community websites including blogs, cafés, and other personal homepages can be harmful websites for children. Furthermore, these blogs and cafés are easier to open and close than classic personal homepages, and if some community site contains harmful information, it is hard to block.

In case of video files, filtering software, X-Keeper and Momi, may also use database to block execution of adult video files. Thus, a video file which is blocked by filtering software can be executed if a video file is encoded in other formats.

As with video file re-encoding, some filtering software cannot block harmful websites if the user uses various web browsers such as Firefox, Chrome, and Opera etc. The reason for this problem is implementation dependency of filtering software. To block websites, filtering software has to obtain URLs from the web browser. If implementation of filtering software is dependent on a specific web browser, it cannot block access to harmful websites when using other web browsers.

Bypassing Filtering Software Functionalities. It is possible to bypass filtering software functionalities. In this section we describe bypassing methods of filtering software functionalities. We suppose that the attacker – who wants to access adult contents by bypassing filtering software – is a computing novice. Table 9 shows bypassing methods of filtering software functionalities.

Table 9. Bypassing Filtering Software

Bypass	*Description*
Modification or Uninstall	The attacker can delete software from a computer, or modify the configuration of a filtering setting, thus software has to be protected using authentication.
Closing Process	The attacker can terminate an execution filtering software process using task management.
Editing Register	The attacker can edit the register to block automatic execution of filtering software when the computer completes booting.
Change of System Time	The attacker can edit the system time to avoid locking of the computer by filtering software.
Using Utility	The attacker can use various utilities to avoid or block execution of filtering software.
Using Keylogger	The attacker can use a keylogger to obtain an administrator password.

In Microsoft Windows, users can manage processes on their computers using task manager, which is a basic utility built into Windows. Task manager in Windows is easy to execute – the user just simultaneously presses the keyboard control key, alter key and delete key, and can terminate processes, thus the attacker can use this tool to deliberately terminate filtering software process in order to avoid filtering adult contents. The purpose of editing the register is to prevent automatic execution of filtering software. Generally, execution of filtering software is part of system booting, because most filtering software is registered in the list of programs the operating system has to execute after booting. Thus, the attacker can edit the register to eliminate filtering software in the list. Most filtering software provides locking of the computer based on the system time. The purpose of this functionality is to limit the computer usage time. However, if an attacker can edit system time, then an attacker can use his/her computer all day. As like task manager, an attacker can use various utilities, which provide more detail system management than task manager, to prevent execution of filtering software. Finally, most filtering software uses a password to administrator authentication. If an attacker obtains a password using a keylogger, he can neutralize filtering software. Table 10 shows the results of avoidance testing for seven types of filtering software.

Table 10. Result of Bypassing Test

S/W	*Modification or Uninstall without authentication*	*Closing Process*	*Editing Register*	*Change of System Time*	*Using Utility*	*Using Keylogger*
ParentalControl	O	X	X	O	O	O
Magic Desktop	X	X	X	X	X	X
Web Filter PC	X	X	O	O	O	O
X-Keeper	X	X	X	X	X	O
Momi	X	X	X	X	X	X

In the table, “O” means that it is possible to bypass filtering software, and “X” means vice versa.

In bypass testing, we use “Click to Tweak Basic version” for “Using Utility” testing. This utility provides system management, system optimization, register optimization, trash file optimization and etc. Furthermore this utility provides a closing process, and it can prevent execution of filtering software.

Especially, Magic Desktop is software for toddlers rather than children, and it provides a new interface for the user, called Magic Desktop. Thus, in case of Magic Desktop, the user only can use the computer in circumstances controlled by Magic Desktop, so it is impossible to access adult contents.

A keylogger is a program that audits all the keyboard and/or mouse input and makes a record. Using a keylogger, the attacker can obtain an administrator password and then, neutralize filtering software. Unfortunately, with the exception of Momi, the five filtering software are vulnerable to keyloggers. A keylogger is not a complex program, thus it is easy to acquire a free keylogger from the web.

5 Improvement of Filtering Software

5.1 Requirements for Filtering Software

In this section, we derive requirements for filtering software based on our analysis. First, filtering software has to block access to adult websites, and particularly, its implementation must be independent of the web browser. Furthermore, filtering software has to adopt a machine learning technique to filtering functionality, thus it has to overcome the limitations of a database. A content-based machine learning technique can determine whether it is harmful or not.

Second, filtering software has to prevent execution of files. Harmful files can be classified into texts, images and videos, and filtering software has to provide contents-based filtering functionality for each type of file.

Third, filtering software has to prevent program execution. The main purpose of this functionality is to block harmful games containing violence, porno or gambling. It is also necessary to control execution of online and offline game. In addition, with this functionality, user also can block execution of messenger, P2P, and e-mail programs.

Forth, to block harmful information via an unknown path, filtering software has to make an audit record for all auditable events such as a list of visited websites, executed files and programs, computer usage (hours per day) and etc. Best way checking unknown path is to make screen shots periodically.

Fifth, filtering software has to provide time limitation in computer usage. Time limitation on computer usage can reduce the probability of contact with harmful contents and the potential of young people to become addicted to the computer.

Sixth, filtering software has to be tolerant for functionality bypassing. As mentioned in section 4, there are various methods to bypass filtering functionality, and filtering software has to overcome these bypassing. Table 11 shows all requirements for filtering software.

Table 11. Requirement for Filtering Software

Requirements	*Description*
R1.Filtering websites	Filtering software has to block access to harmful websites, and particularly, its implementation has to be independent of the web browser. Filtering software must be implemented with contents-based filtering functionality.
R2.Blocking files	Filtering software has to provide functionality of blocking file execution.
R3.Blocking programs	Filtering software has to provide functionality of blocking program execution.
R4.Audit	Filtering software has to make audit records for all auditable events such as a list of visited websites, executed files and programs, computer usage (hours per day). Filtering software has to make screen shots periodically.
R5.Time Limitation	Filtering software has to provide functionality to limit computer time.
R6.Preventing bypassing	Filtering software has to block functionality bypassing.

5.2 Essential Functionalities for Filtering Software

In the previous section, we defined the requirements of filtering software. In this section, we derive the essential functionalities of filtering software based on the requirements. Table 12 shows essential functionalities of filtering software.

Table 12. Essential Functionalities for Filtering Software

Functionality	*Description*
Filtering Websites	Contents-based filtering Independent operation on web browser Filtering based on forbidden words registered by user Filtering based on URL which contains forbidden words registered by user
Blocking Files	Contents-based filtering Text, image, and video file filtering available
Blocking Programs	Providing list of all programs installed on computer Blocking program selected by user
Audit Record	Making screen shots periodically Making audit records for all auditable events such as list of visited websites, executed files and programs, computer usage (hours per day)
Alarm	Alarm to administrator using SMS, e-mail etc.
Protection	Providing authentication to configure or uninstall filtering software
Bypass Elimination	Anti Keylogger Preventing modification of system setting, such as system time, DNS server, registry information
Time Limitation	Locking computer based on timetable

6 Conclusion

The openness of the web allows user to easily get information anytime and anywhere. However, openness has disadvantage also. For example, children can get inappropriate information like pornographic contents in the web. Actually, according to the Korea Communications Standards Commissions, 83.4% of respondents encounter harmful contents on the web.

To block adult contents from children, there are several filtering software. Filtering software is one of the effective methods to block adult contents. General functionalities of filtering software are blocking access to adult contents according to predefined criterion. However, filtering software cannot entirely block harmful information, because it has some limitations.

Thus, in this paper, we introduced several filtering software, and then analyzed their effectiveness. Furthermore, we discussed the limitations of filtering software and proposed their essential requirements based on our analysis.

References

1. Survey of young people using illegal harmful information, Report published by Korea Internet Safety Commission (2007)
2. Survey of young people using illegal harmful information, Report published by Korea Internet Safety Commission (2006)
3. `http://www.wikipedia.org`
4. `http://checked.icra.org/`
5. Kuffner, A., Rebles-Kelly, A.: Image Feature Evaluation for Contents-based Image Retrieval. In: HCSNet Workshop Proceedings: Use of Vision in Human-Computer Interaction, Canberra, Austrailia, vol. 56, pp. 29–33 (2006)
6. Bertino, E., Ferrari, E., Perego, A.: Web and Information Security, pp. 112–116 (2006)
7. `http://www.safenet.ne.kr`
8. Kim, Y., Nam, T., Won, D.: A Distinction Technology for Harmful Web Documents by Rates. In: Korea Information Processing Society, vol. 13C, pp. 859–864 (December 2006)
9. Park, S.B., Tae, Y.S.: Machine Learning and Information Retrieval. Journal of Information and Science 25(3), 5–11 (2007)
10. Zhang, B.T.: Next-Generation Machine Learning Technologies. Journal of Information and Science 25(3), 96–107 (2007)

A Smart Movie Recommendation System

Sang-Ki Ko[1], Sang-Min Choi[1], Hae-Sung Eom[1], Jeong-Won Cha[2], Hyunchul Cho[3], Laehyum Kim[3], and Yo-Sub Han[1,*]

[1] Department of Computer Science, Yonsei University, Seoul, Republic of Korea
{narame7,jerassi,haesung,emmous}@cs.yonsei.ac.kr
[2] Department of Computer Engineering, Changwon National University, Changwon, Republic of Korea
jcha@changwon.ac.kr
[3] Intelligence and Interaction Research Center, KIST, Seoul, Republic of Korea
{hccho,laehyunk}@kist.re.kr

Abstract. We propose a movie recommendation system based on genre correlations. We modify the previous algorithm; we use a list of movies as input instead of genre combinations. We implement a new recommendation algorithm as Android application with additional functions. By combining with existing web services such as Google Movie Showtimes and Open APIs, our system can recommend movies playing in cinemas currently and show the detailed information of movies. Location-based function is also implemented. We utilize GPS information of mobile device and web service provided by Google Maps for recommending suitable cinemas for users with mobile devices.

Keywords: recommendation system, movies, smartphones, Android.

1 Introduction

As information technology develops in these days, we can easily obtain the information from the Internet. Since there are massive materials on the Internet, it is difficult to use all of them efficiently. Thus we should choose which materials to use. This means that we need to know which one is useful and which is not for better recommendation.

Nowadays, smartphones become one of the most important tools for our life. Most of smartphone users tend to use their phones instead of computers when they search information, since mobile phones are more portable and smartphones has many searching applications for various objectives.

We propose a smart movie recommendation system for smartphones. We use a recommendation technique based on the genre correlations investigated by Choi and Han [1]. We implement the proposed system on Android platform.

In Section 2, we revisit the previous research with respect to recommendation systems. Then, we propose our approach that applies recommendation technique to mobile applications for movie recommendation systems in Section 3. Then we present the result of implementations on Android platform in Section 4. We conclude with future works of this research in Section 5.

* Corresponding author.

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 558–566, 2011.

2 Related Works

There are lots of recommendation techniques investigated by many researchers [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]. Recommendation system attempts to recommend information items that are likely to be of interest to the user. There are four kinds of recommendation techniques, content-based, association, demographic and collaborative method.

Content-based method uses item-to-item similarity. If a user like B, we recommend A that is similar to B. Association method also uses item-to-item similarity. In this method, we do not decide whether actually they are similar or not. If items have high correlation with each other, we decide that they are similar. Demographic method and collaborative method use people-to-people similarity both. Demographic method needs actual features of people to decide whether they are similar. Collaborative method uses correlation between users.

Basically, our system uses item-based method. The detailed explanation for our recommending algorithm will be covered later.

2.1 Collaborative Filtering

As the amount of information in the world is increasing very quickly, we need techniques to find relevant information efficiently. One of such technique is to use a recommendation system and the collaborative filtering [7, 8, 9, 10] is one of the most promising methods. Collaborative filtering is a method of making automatic predictions about the interests of a user by collecting taste information from many users. The underlying assumption of this approach is that those who agreed in the past tend to agree again in the future.

We need to build a database of preferences for items by users at first. A new user is matched against the database to discover neighbors that are users who have historically had similar taste to the new user. Items that the neighbors like are then recommended to the new user as he will probably also like them. This approach has been very successful in many recommendation systems.

2.2 Genre Correlation

General collaborative filtering approach is based on user preferences. This implies that the system should wait until it has enough input data from users. Researchers proposed several methods to avoid this problem [1, 2, 3, 4, 5, 6]. One of such approaches is to use information that is reliable and available initially. Notice that we cannot always have such information available. Thus, we choose a movie recommendation system domain since a movie has a category information (called genre) given by experts.

Recently, Choi and Han [1] proposed a movie recommendation system based on genre correlations. Their system does not require lots of user preferences. The system first calculates genre correlations based on the genre combinations of each movie. Then the system applies the genre combination of all movies and user-preferred genres to the average rating of each movie based on the genre correlations. Finally, it ranks movies according to the newly computed point.

3 Our Approach

We propose a movie recommendation system based on genre correlation. Choi and Han [1] suggested the method that users should input their favorite movie genres into the recommendation system manually and the system calculates the recommendation points. On the other hand, our recommendation system uses movie lists as input and obtains the genres of movies in lists and thus the preferences of users. (We profile movie preferences of users.) This step assumes that users would prefer genres appeared in the list to other genres.

3.1 Calculating Genre Correlation Based on Movie List

Since we have a list or several lists of movies for recommendation, we should figure out the number of appearances of genres in the list. If there are two movies with the first movie having comedy and drama as a genre combination and the second movie having romance and drama as a genre combination, the number of appearances of drama is two. In the same way, comedy is one and romance is also one. These numbers are used for weighting each genre when we calculate the recommendation points.

3.2 Movie Information Retrieval

Our application retrieves the movie information from the famous portal sites. When a user tries to see the information of a movie, this application sends the title of movie to

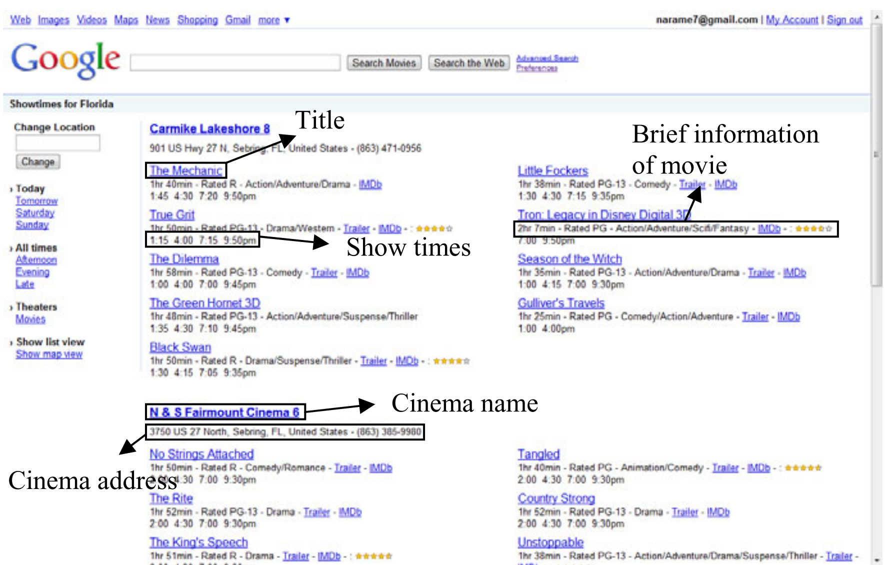


Fig. 1. This page shows the show times of movies in Florida, United States of America. Google Search provides this service via web pages. You can enter the location where you are living instead of Florida. The address of this page is http://www.google.com/movies?near=Florida

Open API services. Then they return the information of movie in XML format. In our application, since we develop this in South Korea, we use Korean portal sites offering Open API services for movie information.

3.3 Location-Based Cinema Recommendation

Google Search offers Movie Showtimes [11] service that provides the show times of movies to users. Users can know at which cinema a certain movie is playing near them and show times of movies and even the address of cinemas. We use this service for our application. Detailed explanation of how we used this web service for our system is described in Section 4.3.

4 Implementation

We implement our recommendation system on Android OS. Testing device is HTC Desire with Android 2.2 version. Since Android platform uses Java programming language, we use Java for developing this application. PHP language is also used for server programming. The overall structure of our system is depicted in Fig. 2.

There exist three objects in our system: server, mobile device and web services. We use a server here because the size of movie database is too large to handle on mobile device. When a mobile phone recommends movies, we need to calculate recommendation points of all movies according to genre correlation matrix and the calculation consumes too much time and resource when it is computed on a small device such as smartphone. Therefore, a mobile device sends the list of movies and the server calculates recommendation points of all movies stored in server and sort them. Finally, the server returns the resulting list, that is, the list of recommended movies to the mobile phone.

In Fig. 2, there are three web services used in our recommendation system. We use Google Movie Showtimes service and Google Maps service. As we briefly cover in Section 3.3., Google Movie Showtimes service provides regional cinema information and movie information. They provide these two kinds of information in http://www.google.com/movies page. We can obtain regional cinema information by putting the address as a GET parameter.

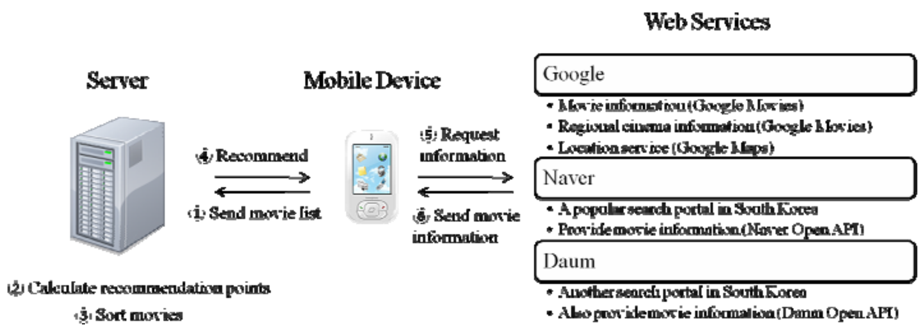


Fig. 2. This diagram describes the procedural steps of our recommendation system. Mobile device generally means a smartphone or a tablet PC using Android OS with GPS information.

Naver[1] is a popular search portal in South Korea, with a market share of over 70%, compared to 2% of Google. They provide Internet services including a news service, an e-mail service an academic thesis search service and so on. They also provide Open API service that we can get movie information by using the title of movie. If we request the movie information with the title of movie, they return the information of movie as XML form. The information includes the English title and Koran title of movie and directors, actors, user rating, related links and so on.

Daum[2] is also a popular web portal in South Korea. Daum offers many Internet services to web users, including a popular free web-based e-mail, messaging service, forums, shopping and news. They also provide Open API service relating to movies.

4.1 Development Environment

We work on Windows 7 64bit OS. Tool used for our development is Eclipse Helios version and JDK version is 1.6.0. We also use a server for server programming. Our server is using CentOS release 5.5 as an operating system and kernel version is 2.6.18. The version of PHP language installed on our server is 5.2.10 and the version of MySQL is 5.0.76.

4.2 GroupLens Database on Our Server

We use an open movie database called *GroupLens* database[3]. The GroupLens database has three sub-databases: movie database, user database and rating database. The movie database has information of 10681 movies. We create a MySQL database on server for storing this database.

4.3 Recommendation Result Based on Movie List

Our system has two kinds of movie recommendation methods. The first method is recommending the movie in the database. Since we use the GroupLens database that has 10681 movies, the recommendation result always consists of the movies in the GroupLens database. The second method is recommending the movie that is now playing in cinemas.

The right picture of Fig. 3 shows the result of recommending the movies that are now playing in cinemas. The English title of the first movie in the list is 'The Private Lives Of Pippa Lee' and the second movie is 'Gulliver's Travels'. The genre of these movies is both comedies. The genres of other movies in the list are all comedy or animation. Note that the genre combinations of three movies in the left list contain comedy and animation. We can confirm that the result of recommendation is quite reasonable.

---

[1] http://www.naver.com
[2] http://www.daum.net
[3] http://www.grouplens.org/

Fig. 3. The left shows the list of movies and the center shows the result of recommendation when we apply our method to the list. The right shows the result of recommending the movies that are now playing in cinemas.

4.4 Open API Services

When our system displays the result of recommendation as a list of movies, users can see the detailed information of the movies. This information retrieved is from Open API services of Naver and Daum [13, 14], famous South Korean portal sites. The process of retrieving information is quite simple. When users want to see the detailed information of the movie, our system sends the name of movie to Open API services. Since sometimes there is no available information of certain movie, we call two Open API services one after the other. That means if there is available information in first Open API service, we do not need to call the other. The retrieved information is displayed on mobile device with Android layout as illustrated in Fig. 4.

Fig. 4. These pictures show the detailed information of two movies 'Toy Story 3' and 'Love Actually'. We capture in Android SDK Virtual Device. Since the retrieved information is from South Korean portal sites, the contents are in Korean. The contents consist of poster, Korean title, English title, directors, actors, user rating, and summary of two movies.

Since the return forms of Open API services are in XML format, we implement Java-based XML parser to display this information on Android layout.

4.5 Location-Based Cinema Recommendation

Google Movie Showtimes [11] service provides the movies which are now playing in cinemas and regional cinema information. We can know the cinema information of certain region by inputting the address or GPS information as a GET parameter.

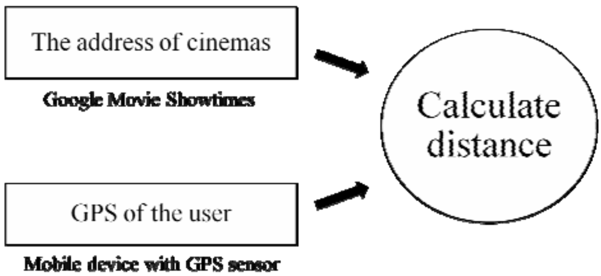


Fig. 5. The process of obtaining the distance between the user with a mobile device and the cinema. GPS of the mobile device can be obtained by GPS sensor in the mobile device.

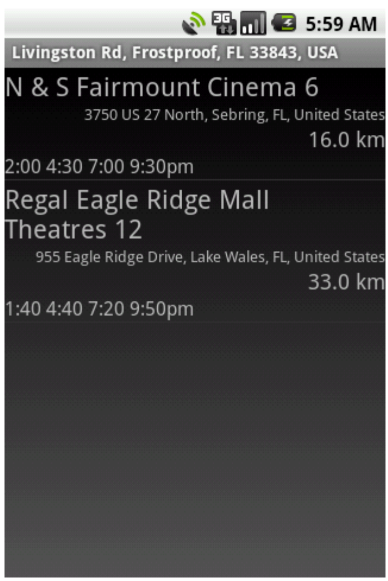


Fig. 6. Our application shows a list of cinemas where certain movie is playing. This picture shows the list of cinemas with some information such as name of cinemas, address of cinemas, the distances between the mobile device and the cinemas and the show times of the movie at each cinema.

Google Maps [12] offers the service that converts GPS information to an address or an address to GPS information. Our system provides the distance between the user and the cinema where the recommended movie is playing. Fig. 5 presents the process of obtaining the distance using Google web services.

Fig. 6 shows a list of cinemas where one of recommended movies is playing. We can find name, address, distance from the user, show times of the movie of each cinema.

As seen in Section 3.3, Google Movie Showtimes service provides their information in HTML web pages. To parse the page and retrieve useful contents for us, we implement HTML parsing program with PHP on our server. When our application on Android device request the cinema information with the title of movie and the current location, the PHP program returns the cinema where the movie is playing and close to the current location of Android device.

5 Conclusions

We propose a movie recommendation system based on genre correlations. In this paper, a list of movies is used for calculating recommendation points instead of genres. We implement the proposed algorithm as a mobile application on Android OS. By combining with existing web services such as Google Movie Showtimes, our application provides a list of recommended movies that are now playing in cinemas and the regional information of cinemas.

We use genre information here for recommending movies in this paper. Since there are more features in movies such as directors, actors and so on, we can utilize those features in our future works.

In future, we can connect our service with ticket reservation system of cinemas. Then, the users can be recommended and can reserve the tickets of cinemas around the users at a time by using our application.

Acknowledgement. This research was supported by the IT R&D program of KME/IITA 2008-S-024-01. Han is supported by the Basic Science Research Program through NRF funded by MEST (2010-0009168).

References

1. Choi, S.-M., Han, Y.-S.: A content recommendation system based on category correlations. In: The Fifth International Multi-Conference on Computing in the Global Information Technology, pp. 1257–1260 (2010)
2. Huang, Z., Chen, H., Zeng, D.D.: Applying associative retrieval techniques to alleviate the sparsity problem in collaborative filtering. ACM Trans. Inf. Syst. 22(1), 116–142 (2004)
3. Ishikawa, M., Géczy, P., Izumi, N., Morita, T., Yamaguchi, T.: Information diffusion approach to cold-start problem. In: Web Intelligence/IAT Workshops, pp. 129–132 (2007)
4. Popescul, A., Ungar, L.H., Pennock, D.M., Lawrence, S.: Probabilistic models for unified collaborative and content-based recommendation in sparse-data environments. In: Proceedings of the 17th Conference in Uncertainty in Artificial Intelligence, pp. 437–444 (2001)

5. Schein, A.I., Popescul, A., Ungar, L.H., Pennock, D.M.: Methods and metrics for cold-start recommendations. In: The 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 253–260 (2002)
6. Wilson, D.C., Smyth, B., O'Sullivan, D.: Sparsity reduction in collaborative recommendation: A case-based approach. IJPRAI 17(5), 863–884 (2003)
7. Hill, W., Stead, L., Rosenstein, M., Furnas, G.: Recommending and Evaluating Choices in a Virtual Community of Use. In: Proceedings of CHI 1995 (1995)
8. Konstan, J., Miller, B., Maltz, D., Herlocker, J., Gordon, L., Riedl, J.: GroupLens: Applying Collaborative Filtering to Usenet News. Communications of the ACM 40(3), 77–87 (1997)
9. Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., Riedl, J.: GroupLens: An Open Architecture for Collaborative Filtering of Netnews. In: Proceedings of CSCW 1994. Chapel Hill, NC (1994)
10. Shardanand, U., Maes, P.: Social Information Filtering: Algorithms for Automating 'Word of Mouth'. In: Proceedings of CHI 1995, Denver, CO (1995)
11. Google Movie Showtimes, http://www.google.com/movies
12. Google Maps, http://maps.google.com/
13. Naver Open API, http://dev.naver.com/openapi
14. Daum Open API, http://apis.daum.net

Interactive Personalization of Ambient Assisted Living Environments

Alexander Marinc, Carsten Stocklöw, Anreas Braun, Carsten Limberger, Cristian Hofmann, and Arjan Kuijper

Fraunhofer Institute for Computer Graphics Research,
Fraunhoferstr. 5, 64283 Darmstadt, Germany
{alexander.marinc,carsten.stockloew,andreas.braun,carsten.limberger,
cristian.hofmann,arjan.kuijper}@igd.fraunhofer.de

Abstract. Ambient Assisted Living (AAL) comprises methods, systems, and services applied to improve the quality of daily life for humans, especially elderly people. Recent research emphasizes the implementation of comprehensive AAL platforms which control all technological components included in the entire environment such as one's apartment. The behavior of the system is often determined by a specific set of rules. Thus, personalization according to the person's needs and preferences includes a configuration of the given rule system. Assuming that configuration is not only conducted by technical staff but also by the person him or herself, this process can be regarded as complex, requiring technical knowledge. In this work, we present an interactive and architectural approach to support at the personalization of an AAL system by different types of users.

Keywords: Ambient Assisted Living, End User Configuration, Personalization.

1 Introduction

Ambient Assisted Living (AAL) comprises methods, technological systems, products, and attendances applied to improve the quality of daily life for humans in different periods of life. Considering predictions of the demographic changes in society, AAL particularly focuses on elderly people. Especially, the integration of technology is user-centered, that is, it is adapted according to the specific needs of assisted persons and unobtrusively embedded into the environment. Recent research directions emphasize the implementation of comprehensive AAL platforms which manage and control all technological components included in an entire environment such as ones apartment. Among these technological components are i/o-devices like interactive TV, RFID tags or capacitive sensors for localization and activity recognition, or actuators such as robotics.

Characteristical AAL platforms bear specific challenges referring to their personalization, regarded as the configuration and modification of the rules which determine the behavior of the environment according the specific needs and preferences of the assisted person. First of all, such challenges arise from the diversity

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 567–576, 2011.

of heterogeneous components and information the platform needs to handle, e.g., computing unit(s), devices, software components, services, or different participants. Second, such environments are usually driven by a specific set of rules which finally form the application tier. Here, rule-based applications define the overall behavior of the system and determine the state of different elements potentially resulting from specific predefined events. As a result, personalization processes require technical knowledge and skills such as specific notations or programming languages, which are normally not among the standard competencies of regular users, and even less of assisted persons. In this context, we must consider various types of users which differ according to their technical knowledge, skills and expertise on the one hand, and their physical capabilities or disabilities [7]: *expert*, *normal*, and *impaired* user. While the first type can particularly be associated with technical staff, normal and impaired users refer to the assisted persons. Here, we regard impaired users as persons which have physical and mental disabilities.

In this paper, we present a new approach for the interactive configuration of comprehensive Ambient Assisted Living Environments on the level of authoring tools, focusing on the application of AAL at home. First, essential requirements are discussed which refer to the given AAL-system and the different types of users. In Section 3, configuration parameters are elucidated, explaining what needs to be configured regarding objects of the AAL space as well as the given system of rules. As central contribution of this work, in Section 4 we present an interactive concept to support different types of users during the personalization process. Here, we specify different methods for multimodal interaction and abstraction of information and access to the system. In Section 5, we introduce architectural components which realize the established concepts, and integrate them into an existing AAL platform developed in the scope of the *UniversAAL*[1] project. Concluding, we summarize the presented results and describe our future work.

2 Requirements

In this section, first we give an overview of the general requirements which have to be fulfilled by a platform to allow an efficient personalization of a system. Then we present requirements that are of interest for the different user-types such as defined in Section 1.

2.1 General Considerations

To allow an intuitive and easy interaction, a technical system needs to fulfill a minimal set of requirements regarding the hard- and software structure. According to our research focus, these requirements generally depend on a) the number of users that are using the system and b) the users' knowledge and skills.

Handling multi-user-systems is a part of current research. The limitations in this case are not primarily given by the software but by sensor systems which track and distinguish between different persons within the same living area. Recent approaches include the equipment of every user with small active sensors or passive tags that can be detected by ultrasound or radio signals, or vision based methods with time-of-flight or video cameras. However, these methods require the user to either wear sensors/tags or suffer from user acceptance in case of cameras. Due to these limitations and assuming that a majority of the elderly people is living alone at home [3] we will handle in this work basically one-user-systems.

For a user interface, considering the user's knowledge and skills is an important aspect. The requirements regarding the categories of users are given in Section 2.3.

2.2 Basic Requirements

From a high-level point of view, there is a number of requirements that a platform needs to fulfill to enable a system to be configurable by a user. For a detailed description, the reader is referred to [8].

R1 - Hardware abstraction layer: A platform for personalization needs to abstract from the hardware in a way a user can query all available devices, their functionality and benefit in the overall system, as well as their logical connection.

R2 - Interaction framework: The platform needs to support a framework for managing the interaction between human and machine. This includes the processing of multimodal in- and outputs that enables a user to interact in a way that is most natural to him/her and fits his/her skills and physical needs.

R3 - Rule based system: To realize this, we use a rule-based system to describe behavior in a way that it can be understood and saved by the system, and edited and parameterized by the user to fit it to his needs.

R4 - Service based infrastructure: Services are an adequate choice to create an open system. This allows an understandable description, abstraction, and connection of different distributed functionalities.

R5 - Context reasoning: The intended behavior strongly depend on the current state of the system and the context of the user (e.g. the user is now in the kitchen). Therefore a system is needed that allows to save, read and change the current context as well as making it transparent to the user.

R6 - Semantic descriptions: Semantic descriptions enables discovering and processing of services, and providing contextual data in a way that end users can work with it.

2.3 Requirements for Different Kind of Users

In the following, the requirements for the different user types are defined. These are requirements for the tools that realize personalization on top of the requirements described in the previous section. Some requirements are relevant for all kinds of users, but we assign requirement to users where it is most appropriate.

Requirements for expert users: An expert user has profound knowledge about the system and its configuration. Thus, the rules and parameters could even be shown directly, only small abstraction can be considered helpful. However, the following additional requirements should be fulfilled:

R7 - Direct interfaces: Since an expert user is able to program even complex rules by himself he needs to have direct access to all relevant components.

R8 - System details: An expert user will perform personalization as a daily routine and has detailed knowledge about the technical background. Thus, all details must be available. Rules and parameters can be edited "by hand".

R9 - Direct Feedback: All feedback from the system has to be shown directly to the user without modification or delay.

Requirements for regular users: A regular user is typically not interested in all details about the used technologies of the platform, but he is well trained in using common interaction devices. This kind of user is hard to deal with, because he knows what is possible, but not how to realize it. A personalization tool for a regular user therefore needs to hide technical aspects, but in best case offer the same functionality as for the expert user. The following additional requirements are defined:

R10 - Common user interfaces: For regular users tools for personalization must make use of different interaction techniques with the system. Since they are not interested in details, other modalities like speech are very relevant here to gain a bigger benefit in usage.

R11 - Help-files: A regular user is able to adapt new concepts based on common knowledge. Technical details should be hidden, but hints about them have to be available in tutorials and help-files.

R12 - Up-to-date look and feel: For the acceptance of a system it is particularly for the regular users important to have a modern look and feel of the visual components.

Requirements for impaired users: Impaired users need to be assisted much more by the tools than other users. They are typically not able to use the full functionality given by the system and accordingly they will be offered only basic elements of configuration. An interesting aspect is the automatic learning of the user's behavior to try to estimate his/her current needs. This way, it will be possible to guide him/her to use the system more use- and trustfully. For the impaired user we define the additional requirements in the following:

R13 - Limited access to details: Complex and vital details of the system have to be hidden. This way, the user does not get overwhelmed by the user interface or even trapped in nested dialogs. Additionally, some parameters like emergency processing can not be made inoperative by the user.

R14 - Simulation: Offer the possibility to simulate recently created rules. This means, testing without changing anything in the real system. This can take the fear to "play"with the system.

R15 - Self learning: In particular for the group of impaired users it can be helpful if not only the user has the active role. If the system detects common patterns in the behavior of the user (e.g. switch on TV every morning at 8am), it should suggest to the user the permanent acquisition of automatically created rules.

R16 - System interruption: Impaired users have in most cases no idea how the system is working, therefore it is important that every behavior introduced by the system can be interrupted by the user as easy as possible. This way, it is ensured that a user does not feel patronized by the system.

3 Configuration Parameters

In this section, we provide a more detailed look on the technical elements and properties which are to be personalized. As already stated in Section 2, we assume to have a given Service-Oriented Architecture (Requirement 4). The very basic elements are Service-Descriptions, Service-Calls, and appropriate Service-Responses. In addition to this, we need to have a sort of Reasoning Managing System which is able to create, distribute and save changes in the context of the user (context-events). Finally, the platform must support input and output channels which allow a user to interact with the system (see also Section 5).

3.1 Configuration of Single Elements

One essential object is an editable user-profile which contains general information such as the person's name or physical limitations. The other basic element in the system are the Service-Descriptions that have to be given in same description language. A common way is to use the XML based language WSDL[1], but also semantic expressions can be used, e.g., provided by OWL[2]. In order to make all services available to the user, we have to ensure that enough meta-information is included in order to distinguish between different user types. Additionally, this information must be editable in order to adapt the descriptions according to the individual user's needs (e.g. for a description of a TV "X8300"vs. "TV Living Room"). The needed inputs for Service-Requests can be retrieved by simple questionnaires and the presentation of the Service-Responses strongly depends on the user-profile (e.g. using loudspeakers if the user is blind). The current context of the user (except of his profile) does not have to be configured by the person himself. We must assume that the platform provides an appropriate overview about the available context-events. Furthermore, a prerequisite is that programmers defines suitable names. This gives a user the freedom to include event-reasoning in his personalized rules. Output devices can be included as services and Input-Events are represented as events that provide input for Service-Calls.

3.2 Configuration of the Dependencies between Elements

Besides the configuration of single elements, the main part of personalization is to take control over the system by creating rules that are usable by the platform, combining given single elements. Combining means in the simplest case just to

[1] Web Services Description Language: XML-based language for the Web Services.
[2] Web Ontology Language: RDF/XML-based language for the Semantic Web.

group services or events. Multiple services are called at the same time and will be executed after all required inputs of all services in the group are given by the user. Examples for sophisticated methods such as entire workflows between services. This offers the possibility to define restrictions in the process (e.g., take service X if Y failed and Z otherwise). Since many interactions are complex and multimodal it most be also possible to combine events (e.g. combine “switch lamp on”and “user is in living room”to create the event “switch lamp in living room on”) for futher use. The most important aspect is the combination of services and events. This is what offers the possibility to directly control the behavior of the system. If a (maybe combined) event like an explicit command is triggered, an associated service (this again can consist of many services) should be triggered.

3.3 Representation of Rules

All rules and combined services have to be represented in an appropriate way in the system. To realize this, different forms of representation have been researched in the past, e.g. for the definition of workflows the description languages BPEL[3] or XPDL[4] can be used. The platform itself is not bound to a specific representation; there can be multiple components responsible for storing, executing and processing the various types of description languages. In this case, it must be possible to access and configure the rules contained in these components to provide the possibility to combine them into higher level rules, and to provide appropriate interfaces to present them to the user.

4 Interaction

In this section, we present an interaction concept which supports the process of customizing the system of rules of an AAL-system. The main idea can be exemplified with the simple scenario “evening activity”. *After dinner, the assisted person usually reads a book in the living room. For this purpose, when he or she leaves the dining room the lights are automatically switched off. In the living room, the ceiling lightings are dimmed and a reading lamp positioned next to the couch is switched on. In the course of time, the person's ability to see has decreased. As a consequence, instead of reading, the person is going to listen to the local radio program. Hence, the reading lamp remains off, and the hifi-system is turned on with the preset radio channel.* Here, the underlying rules system needs to be modified. As already mentioned, these changes can be conducted by different kinds of users with different knowledge and skills. In order to take this aspect into consideration, we enable a multimodal communication with the system on the one hand, and define varying abstraction levels on the other hand.

[3] Business Process Execution Language.
[4] XML Process Definition Language.

4.1 Multimodal Interaction

Going back to the “evening activity”scenario: assuming that the person’s ability to see has decreased, this not only results in a change of the activity, but has also an impact on the way a respective modification is conducted. In this example, a “traditional” interaction by means of a *Graphical* User Interface (GUI) would certainly be the wrong choice. A better solution would be to allow a more intuitive and accustomed interaction, for instance, by pointing at the reading lamp and expressing the command “switch off” in a verbal way. If the person repeats this on several evenings, the system would change the related rule after a confirmation request. Thus, we introduce three different modalities which can be combined in the scope of one single interaction, *GUI-based*, *speech-based*, and *gesture-based* (see Fig. 1).

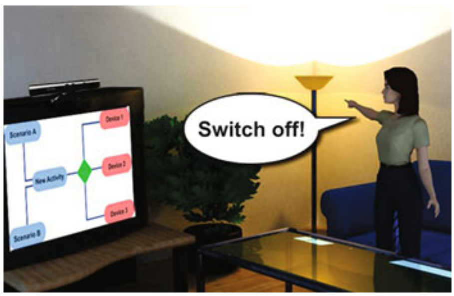


Fig. 1. Multimodal personalization by means of a GUI, speech, and pointing

GUI-based. Interaction based on a GUI requires an appropriate visualization of the given rules system. Thus, all building blocks of the system have to be graphically presented to the user, supporting the creation of a correct mental representation [5]. Examples for common standard visualization models are activity graphs, workflow graphs, or semantic visualizations [4,5]. Here, configuration is usually performed by Direct Object Manipulation [6].

Speech-based. The person can explicitly communicate with the system in a habitual verbal manner based on an *automatic speech recognition* approach (ASR) [2]. Here, spoken words are automatically converted to text and compared to a repository of predefined commands.

Gesture-based. Also regarded as habitual, gesture-based interaction is based on automatic gesture recognition. In our work, a particular focus lies on the recognition of pointing activities. These can be combined with a spoken command, but also with a graphical representation, e.g., by highlighting the graphical representative of the referenced object.

4.2 Levels of Abstraction

In addition to the options to interact with the system by means of different modalities, we introduce different levels of abstraction in order to meet the needs of persons with differing technical and physical prerequisites. Abstraction implies the presentation of information as well as the access to this data in order to perform modification on the set of rules. Thus, abstraction relates to the visualized information with respect to the amount and/or the level-of-detail. Furthermore, it determines which and how many interactions are allowed out of a set of existing options. In the case of GUI-based personalization, different entities can be presented (see Section 3). For persons with high technical knowledge, the presentation can correspond to the code level, e.g., a workflow definition or an RDF[5] file. In the next level, graphical representatives can be incorporated, for example as part of an activity or workflow graph. Here, a further abstraction can be defined in order to show only basic building blocks of the rules system with a set of few basic graphical elements. With regard to verbal commands based on speech recognition, abstraction does not refer to the presented information, but to the set of commands which can be expressed by the user. That is, different vocabularies or parts of a common vocabulary are assigned to a specific user account. These vocabularies range from the explicit editing of code-level constructs such as RDF tags or SQL queries to simple commands as described in the example scenario. Since this work addresses only pointing gestures, no abstraction levels are defined in the context of gesture-based personalization.

5 Architecture

To support the concepts derived in the previous sections, a generic architecture is presented. This architecture is based on the EU project UniversAAL, as it is supposed to become a standardized general-purpose platform for AAL-spaces. UniversAAL is a consolidated combination of prior work, not following a completely new approach but rather integrating approved concepts from a variety of projects in this area. UniversAAL already has some of the components necessary to realize interactive personalization. The important parts together with some enhancements and extensions are illustrated in Fig. 2.

The middleware is the only component necessary on ever node in the network. It hides distribution and heterogeneity, and facilitates communication by providing a bus-based system with four buses. The input and output bus handle explicit interaction with the end user while context and service bus realize push and pull mechanisms, respectively, for interoperability of various components at a semantic level using ontologies and technologies like RDF and OWL[6].

The remaining components are of particular interest and are described in the following subsections.

5.1 Multimodal Interaction

The interaction with the user is probably the most important aspect of a platform to be fulfilled for user acceptance. According to the huge variety of preferences,

[5] Resource Description Framework.

[6] Web Ontology Language.

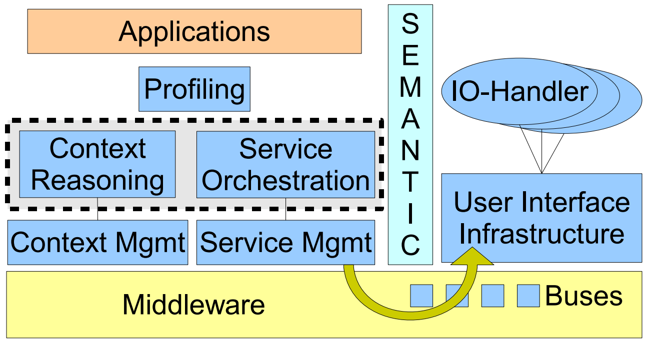


Fig. 2. Architecture of the system, orchestration and reasoning (highlighted by a dotted box) are the parts to be personalized

skills, and impairments, the platform should support multimodal interaction and integrate different modalities. The architecture supports this by integration of various so-called IO-handlers[7] to support different in- and output devices and combinations thereof for interaction with the end users. One IO-handler is responsible for a defined set of modalities and deals with modality fusion and fission. Thus, a different IO-handler can be used for the respective end user. Which handler is used, can be determined by the profiling component, that provides a user model for handling user identity, capabilities, constraints and preferences. Here, a distinction between different user types and their preferred interaction method can be realized.

5.2 Context Reasoning and Service Orchestration

Complex services may not be resolved directly, but by intelligent strategy planning and composition of simpler services. The composed service is then available and registered at the service management component just like regular services.

The context management is responsible for all data and events that can be shared within the system based on a set of shared models and guarantees a certain level of persistence to reflect the current state and allows for querying past events when needed. This data is then used for a rule-based reasoning to derive higher-level situational data from low-level sensor events.

Service orchestration with parameterization and workflow execution as well as the rules for context reasoning are in this work the parts that can be personalized.

6 Summary

In this paper, we present a new approach for the interactive personalization of comprehensive Ambient Assisted Living Environments. This approach includes

[7] Detailed information about the general UI Framework at a technical level can be found in [7].

the specification of essential system- and user-centered requirements, as well as configuration parameters which are to be modified. Moreover, we present an interactive concept to support users during the personalization process. Here, we establish methods for multimodal interaction and abstraction of information and access to the system in order to consider different types of users. Furthermore, we define architectural components which realize different aspects of the presented concepts. As a result, we expect a significant improvement regarding the understanding and the execution of personalization processes. As part of our future work, this improvement will be investigated in the scope of user studies based on a respective prototype implementation. Furthermore, additional interaction modalities will be integrated.

Acknowledgements

This work is partially financed by the European Commission under the FP7 IST Project *UniversAAL*, under grand agreement FP7-247950.

References

1. UNIVERsal open platform and reference Specification for Ambient Assisted Living, http://www.universaal.org/ (accessed February 2011)
2. Baker, J., Deng, L., Glass, J., Khudanpur, S., hui Lee, C., Morgan, N., O'Shaughnessy, D.: Developments and directions in speech recognition and understanding, part 1 (dsp education). IEEE Signal Processing Magazine 26(3), 75–80 (2009)
3. Meyer, S., Mollenkopf, H.: AAL in der alternden Gesellschaft - Anforderungen, Akzeptanz und Perspektiven. VDE Verlag Gmbh (2010)
4. Nazemi, K., Ullmann, T.D., Hornung, C.: Engineering user centered interaction systems for semantic visualizations. In: Stephanidis, C. (ed.) UAHCI 2009. LNCS, vol. 5614, pp. 126–134. Springer, Heidelberg (2009)
5. Rinderle, S., Bobrik, R., Reichert, M., Bauer, T.: Business process visualization - use cases, challenges, solutions. In: ICEIS 2006 - Proceedings of the Eighth International Conference on Enterprise Information Systems: Databases and Information Systems Integration, Paphos, Cyprus, May 23-27, pp. 204–211 (2006)
6. Shneiderman, B.: Direct manipulation for comprehensible, predictable and controllable user interfaces. In: Proceedings of the 2nd International Conference on Intelligent User Interfaces, IUI 1997, pp. 33–39. ACM, New York (1997)
7. Tazari, M.R.: An open distributed framework for adaptive user interaction in ambient intelligence. In: Proceedings of the First International Joint Conference on Ambient Intelligence, AmI 2010, pp. 227–238. Springer, Heidelberg (2010)
8. Tazari, M.R. (ed.) : AAL Reference Architecture requirements. UniversAAL Deliverable D1.2 (2010)

Development of a System for Proactive Information Service

Myon-Woong Park[1], Soo-Hong Lee[2], Young-Tae Sohn[1], Jae Kwan Kim[1], Ilju Bae[1], and Jae-Kwon Lim[2],

[1] Intelligent and Interaction Research Center, Korea Institute of Science and Technology, P.O. BOX 131, Cheongryang, Seoul, Korea
{myon,ytsohn,kimjk,iljusado}@kist.re.kr
[2] Dept. Of Mechanical Engineering, Yonsei University, 50, Yonsei-ro, Seodaemun-gu, Seoul, Korea
{shlee,ljk1225}@yonsei.ac.kr

Abstract. This paper describes a Context-based Information Service Agent which has been developed in the domain of sports, especially for baseball game which often called a sport based on data. For the implementation of the agent, a knowledge model structure which can define the change of context along the axis of time has been suggested. The usability of the agent has been proved as the service system implemented on mobile phone proactively recommended the observation points and information suitable for the current context of watching the game in non-invasive manner.

Keywords: information service, context-awareness, agent, knowledge model, mobile platform.

1 Introduction

Rapid and remarkable progress in computing environment such as world wide web and powerful yet affordable server enable the application programs to provide humans with more intelligent and proactive information service. Software agent is usually the kernel of the service, and intelligent agents for information service in many different domains such as finance, traffic, e-commerce, etc. have been developed for satisfying user's requirement and providing practical service [1, 2, 3, 4]. Among those agents, the context based information service agent is supposed to recognize current situation or change of situation of information user in order to supply information most useful for the user at the context. Recognition of the situation is based on the overall assessment about all the related factors. For the representation of situation, the change over time should be considered as one of the important factors since the agent needs to adapt to dynamically changing situation and correct its behavior [5]. Most of the existing agent are not that much adaptive to the change over time, though.

The method of knowledge modeling is necessary for expanding the range of representation of situation in order to reflect the change of situation over time. While this paper introduces a case of the implementation of a context aware information agent, a structure of the ontology as a knowledge model which can define the change of situation over time is suggested.

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 577–584, 2011.

Baseball game continues couple of hours and the situation consisted of many factors like pitcher, batter, runner, inning, out count, ball count, etc. keep changing. At every change, game audience might be interested in specific data through which the progress of the game is predicted. For instance, the record on the stolen base of a runner might be more interesting than the hitting average of a batter if the runner is on the first base with out-count one at late inning while both team scored nil. The possibility of stealing base at the circumstance is usually reminded and the record of the runner is provided by commentator in case of broadcasting. There can be many ‘observation points’ such as the possibility of stealing base during the baseball game watch. The above mentioned agent was regarded as the kernel of artifact being capable of playing the role of commentator and recommending proactively the observation points suitable for the game context. The agent has been implemented on a smart phone, which becomes the powerful platform for mobile internet and the most useful personal assistant.

2 Context Based Information Service Agent

As the agent needs to recognize current situation and recommend information appropriate for the context, it requires number of functionalities such as context-awareness, inference, information retrieval, and priority indexing (Fig. 1) [6].

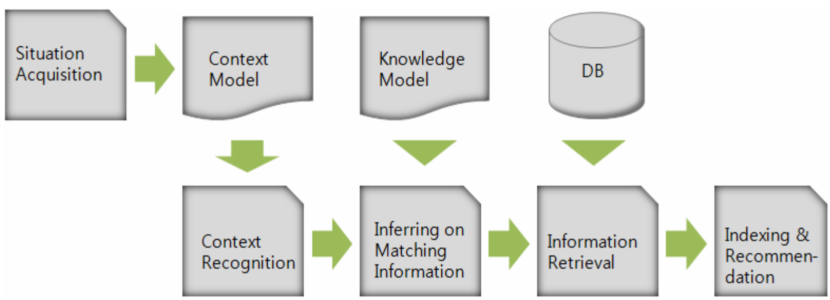


Fig. 1. Functions required for the intended Information Service

For the context awareness, the context modeling technology for the formalization of the recognized context data is necessary. The context recognized through context sensing technology is represented in a formalized model and language. The function to recognize and assess current context by analyzing the model is identified as context recognition. Clear definition on the context elements is necessary for context awareness. And also, grouping of the elements and converting them into knowledge are necessary. Once current context is recognized, the agent refers to the knowledge model and infers the information appropriate for the context. The knowledge model defines context element and information element for generating the relation between two elements in knowledge map. Then, the agent retrieves the specific information relevant to the context from the database and supplies it to the user.

3 Information Service for Watching Baseball Game

There are various records in baseball game as is often referred as the game of data. They are usually provided by caster or commentator. But, inevitably the service is uni-directional and limited in quality and quantity in many cases. Introduction of the 'observation point' and providing the observation point and relevant statistics in suitable timing proactively by the system would improve the service and add the zest to watching the game. The observation point means the group of the player record relevant to the current context, which attracts audience's interest. The system structure of the context-based information agent applied to the baseball game watch is depicted in Fig. 2. The role of each module on the structure diagram is to be introduced, and the flow from context recognition to providing observation points is explained in this section.

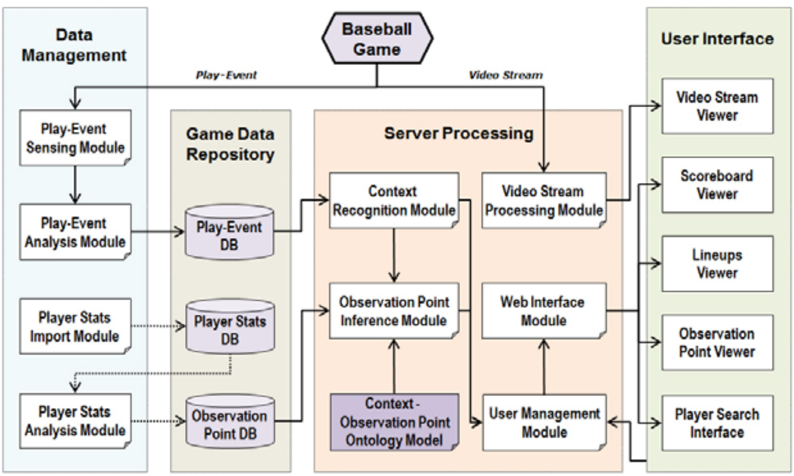


Fig. 2. System Structure of the context-based information agent

Fig. 3. Record Interface

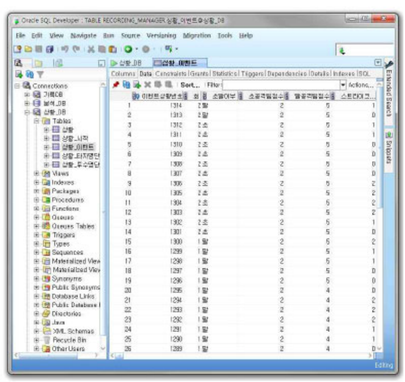


Fig. 4. Context Model

Provisioning of observation points starts when an event occurs, which mean the change of the situation. The event is recorded through the interface as shown in Fig. 3 and analyzed for creating the context model which is then stored in DB as shown in Fig 4. The occurrence of an event is started at the moment of ball drawing by pitcher in this work.

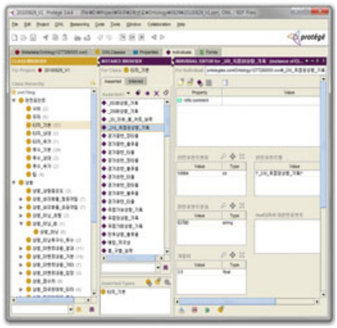


Fig. 5. Modeling of the Relation between Context and Observation Point using Protégé

```xml
<?xml version="1.0" encoding="utf-8" standalone="yes" ?>
<KBD_observationPoint>
  <관전포인트 rank="1" weight="265" no="10064" title="타자_2사_득점권상황_기록" detail="|2사 득점권상
   황 기록#조영훈선수 2사 득점권 상황 기록#22타석, 17타수, 2안타, 0홈런, 2타점,#타율 : 0.118, 출루율 :
   0.318, 장타율 : 0.177#시즌 기록#타율 : 0.299, 출루율 : 0.373, 장타율 : 0.521" />
  <관전포인트 rank="2" weight="258" no="30017" title="투수_시즌달성가능기록_탈삼진" detail="|투수 시즌
   달성가능기록(탈삼진)#김광현선수 시즌 148탈삼진#이번경기 150탈삼진 달성 가능" />
  <관전포인트 rank="3" weight="250" no="10050" title="경기후반_타율" detail="|경기후반 타율#조영훈선수
   경기 후반 타격감 나쁨#경기 후반 타율 : 0.188#시즌 타율 : 0.299#경기 후반 상황 (7~9회) 48타수, 9안
   타, 1홈런" />
  <관전포인트 rank="4" weight="223" no="10083" title="요일별_장타율_화" detail="|요일별 장타율#조영훈
   선수 화요일 장타율 좋음#요일별 장타율 비교#화 : 0.700, 수 : 0.688, 목 : 0.478, 금 : 0.389, 토 :
   0.364, 일 : 0.556#시즌 장타율 : 0.521#화요일 20타수, 6안타, #2루타:2, 3루타:0, 홈런:2" />
  <관전포인트 rank="5" weight="166" no="10061" title="상황중요도_매우높음_상황_기록" detail="|상황중요
   도(매우높음) 상황 기록#현재의 상황중요도(매우높음, 상황중요도 = 3.92)#본 경기의 승패를 좌우할 수 있
   는 중요한 순간#조영훈선수 상황중요도(매우높음) 상황 시 기록#11타석, 8타수, 1안타, 1타점, 0홈런,#타
   율 : 0.125, 출루율 : 0.300, 장타율 : 0.125#시즌 기록#타율 : 0.299, 출루율 : 0.373, 장타율 : 0.521" />
  <관전포인트 rank="6" weight="122" no="10057" title="득점권상황_기록" detail="|득점권상황 기록#조영훈
   선수 득점권 상황에 약함#시즌 득점권 상황 기록#45타석, 34타수, 5안타, 9타점, 0홈런, 타율 : 0.147#시
   즌 기록#137타석, 117타수, 35안타, 16타점, 6홈런, 타율 : 0.299" />
  <관전포인트 rank="7" weight="99" no="10053" title="경기후반_출루율" detail="|경기후반 출루율#조영훈선
   수 경기 후반 출루율 나쁨#경기 후반 출루율 : 0.259#시즌 출루율 : 0.373#경기 후반 상황 (7~9회) 56타
   석, 48타수, 9안타, 5볼넷, 0데드볼" />
</KBD_observationPoint>
```

Fig. 6. Observation Points converted to XML file

The server regularly checks the change in context DB. In case of any change detected, the server identifies the context and extracts observation points relevant to the context. Then, the observation points are served through the user interface. The extraction of the observation points are carried out through the ontology model for the relation between context and observation point. The server compares the current context and the ones defined in the ontology model by parsing the relation model in order to extract the matching context instance. Then, the observation points having the relation with the extracted context are derived. The matching context instance is extracted and the observation point instances having the relation with the extracted context are inferred. The observation point instance includes name of the related table, name of column, searching constraints, etc., with which matching observation points stored in the analysis DB are extracted at the server (Fig. 5). The observation point implies the meaningful data of the players resulted from the analysis on the record DB, which is stored in analysis DB. Weight value is given to each instance at the ontology model in order for comparing the importance of the observation point. The selected observation points adequate for the current context are listed in the order of weight values given to the observation points.

The derived observation points are changed into XML files as shown on Fig. 6 at the web server and sent to UI through http request communication. Through this integrated process, the agent proactively provides human users with context relevant observation points, and the users are able to get interesting record information.

4 Knowledge Model

The structure of the ontology model for the representation of the relation between context and observation point is shown in Fig. 7. Ontology is a knowledge model represented in the form which can be handled by computer. As the definition of all the terms and their relationship are logically described in otology, the model enables formal representation and declarative description of knowledge. Ontology can be applied in many different domains for establishing knowledge base as knowledge management and expansion and inference are possible through ontology.

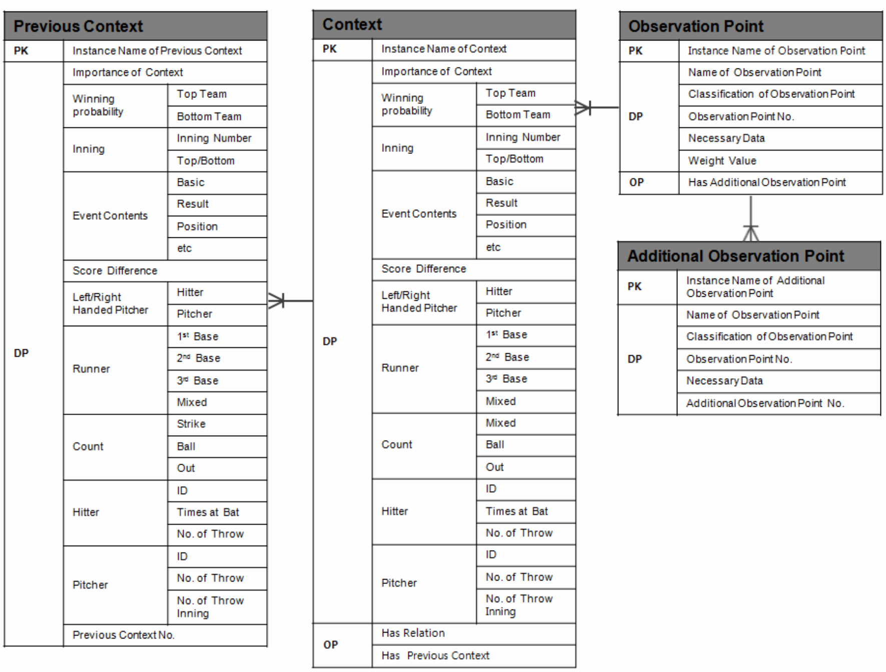


Fig. 7. Model for Relation between Context and Observation Point

Context elements are grouped in ontology structure for the contexts to be recognized according to the data type property. Inning, runner, count, scoring margin, etc. are groups in high level. Inning has inning number and first/second half as lower grouping elements. Count has ball, strike, and out. The ontology structure has been modeled in the way contexts can be specifically represented with these grouping elements. The context formulated with a series of events can be represented since the context class in this study contains the relation with the previous context. The change in context along with the progress of events can be expressed as the sequence is controlled with 'previous context number'. Server can identify the context change over time according to the representation defined in the ontology model. Variety of the observation points have been derived since the context change over time can be expressed and the server is able to recognize the context change through the model represented in ontology.

5 Implementation of Context-Based Information Service for Baseball Game

The information service system for providing observation points and relevant information during the baseball play has been realized in order to assess the usability of the context based information agent. The system has been implemented in the server/client structure based on internet as shown on Fig. 8. The real time information on a game in progress and the video stream of the game are supplied to the server at the ball park. The situation related observation points are derived and indexed by the agent according to the level of relevance. The observation points and related data are stored along with the concurrent video image, and these are supplied to the client on request. iPhone has been used as the client device for mobile service.

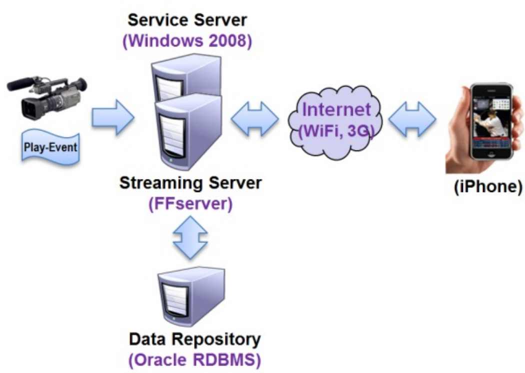


Fig. 8. Platform for the Information Service System

A client App has been developed, which can supply video stream of the ball game, real time observation points, game progress information, player information, etc through the user interface as shown on Fig. 9. The client App regularly requests the server for the observation points along with the timing of the video streaming of a

Fig. 9. User Interfaces for Baseball Information Service

game. The game situation information and the player information are supplied only on demand from user. Only top three of the suggested observation points are displayed at the bottom considering screen size of smart phone. Once any of the observation points is selected by the user, the relevant data are displayed on the top of the screen. The information is overlaid on the video streaming of the ball game. The system has been applied to the baseball games of Korean Series 2010. The system was appreciated that understanding over the progress of the match had been enhanced and the interest and immersion in the game had been increased by the information service.

6 Discussion

In this work, an intelligent information service system for baseball has been implemented as a context based information service agent. Recording interface for the acquisition of the progress information of ball game and for the context modeling has also been designed. An ontology model has been structured for being aware of contextual change over time. The ontology is also designed for parsing data to knowledge and extracting observation points. The observation points which is the representation of groups of the information appealing to the user at a certain context are listed in order according to the relative importance and displayed on the mobile device as shown in Fig. 9.

The implemented system satisfactorily offered information at every change of the situation, i.e. ball count, out count, inning, score and player while the system was applied to the real match of Korean Baseball League. In an actual situation when runners on the first and third base at the out count one for instance, the system suggested the possibility of hit, base steal, and double play. Once the possibility of double play was selected, the record of the batter on the double play against the pitcher during the season and with some other condition was displayed. As the observation points were listed first for the user to select and the record information in detail was supplied only when any of the points was selected, the service was appreciated non-invasive and timely. The scope of this agent can be expanded to other domains like tourism, e-learning, and e-health since the ontology can be modified and expanded as required.

References

1. Dey, A.K., Abowd, G.D., Salber, D.: Conceptual Framework and a Toolkit for Supporting the Rigid Prototyping of Context-Aware Applications. Human-Computer Interaction 16, 97–106 (2001)
2. Kim, J., Hahn, D., Lim, J.: Intelligent Location Guide System based on Context Awareness. In: Proceeding of Conference of Korean Internet Information Society, pp. 683–686 (2010) (in Korean)
3. Ganzha, M., Paprzycki, M., Gawinecki, M., Szymczak, M., Frackowiak, G., Bădică, C., Popescu, E., Park, M.-W.: Adaptive Information Provisioning in an Agent-Based Virtual Organization - Preliminary Considerations. In: Negru, V., et al. (eds.) Proceedings of the SYNASC Conference, pp. 235–241. IEEE CS Press, Los Alamitos (2007)
4. Chang, H., Roh, J., Cho, S.: Context based Use Control Model for Mobile Device. Society of Information Science Journal 14, 63–70 (2008)
5. Dey, A.K., Abowd, G.D.: Towards a Better Understanding of Context-Awareness. In: CHI 2000 Workshop on The What, Who, Where, When, and How of Context-Awareness (2000)
6. Rhodes, B., Maes, P.: Just-in-time information retrieval agents. IBM Systems Journal Special Issue on the MIT Media Laboratory, 39(3and4), 685–704 (2000)

My Personal User Interface: A Semantic User-Centric Approach to Manage and Share User Information

Till Plumbaum[1], Katja Schulz[2], Martin Kurze[2], and Sahin Albayrak[1]

[1] Technische Universität Berlin, DAI-Labor, Berlin, Germany
[2] Deutsche Telekom Laboratories, Berlin, Germany
{Till.Plumbaum,Sahin.Albayrak}@dai-labor.de,
{Katja.Schulz02,Martin.Kurze}@telekom.de

Abstract. With the growing impact of the Web 2.0 on our every day life, people start to use more and more different web based services like Facebook or Twitter. Thereby, they generate and distribute personal and social information such as interests, preferences and goals that are stored in a user profile. This leads to open challenges regarding the users ability to keep track of their personal information but it also offers chances to use this data to enhance personalization and recommendations of existing services. This paper presents a user-centric, thus privacy preserving, system to cope with these challenges and a personal user interface (UI) that allows users to manage and share personal data.

Keywords: information visualization, recommender systems, data mining, user modeling, user-centered design, ontological engineering, semantic mapping.

1 Introduction

With the growing impact of the Web 2.0 on our every day life, people start to use more and more different web based services like Facebook[1], Twitter[2], Flickr[3] or blogs. They use those services to express their opinion, communicate with others and share pictures with friends. Thereby, they generate and distribute personal and social information like interests, social contacts, preferences and personal goals [1]. This user information is usually stored in a user profile deeply buried within every service, only accessible through the User Interface (UI) or API. This affects the users ability to keep track of their personal information. They loose overview what information is stored where and what is public and what private, which leads to open privacy and security challenges. Users who have no overview of the data stored cannot control what data is publicly available and thus, information can be shared by accident. However, the personal information distributed over different services represents an untapped store of knowledge that could be used to enhance personalization and recommendation for existing services. In this work, we present a user-centric, thus privacy preserving, system that

[1] http://www.facebook.com
[2] http://twitter.com
[3] http://www.flickr.com/

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 585–593, 2011.

consists of a semantic layer to aggregate user profiles and a personal user interface to visualize the profile information. The semantic layer aggregates user profiles from different web applications to allow access to information in different user profiles in a unified way. We also present a UI that allows users to keep track of personal information stored in different applications. This helps users to control their personal data and thus it helps to prevent unintended data sharing. The paper is structured as follows: We first give an overview of related work in the field of user model aggregation and sharing. We then describe in detail our semantic approach to aggregate user models and how to access this information in a private and secure way. Finally, we show a system that uses the aggregated information to give users an overview about the personal information stored in different applications.

2 State of the Art

User information is typically stored in proprietary formats defined by each application. Thus, to give users a holistic view on their data, we need mechanisms to aggregate different user profiles. These aggregated user profiles have to be presented in a unified way to have an inter-application understanding of the stored information [2]. Such aggregated user profiles are also the basis for personalization of applications and recommendations [3]. In the research fields of user modeling and user model aggregation, different approaches have been proposed to address the problem of user model heterogeneity and aggregation. These approaches can be categorized into two types [4]:

– Standardization of user models: A common and shared user model standard for all applications.
– User model mediation: A set of techniques to transform or convert data from one user model to another format. This is a practical approach to solve the problem of heterogeneity and allows the aggregation of different models.

The standardization approach has a long research history starting with early works of simple user modeling shells [5] to more sophisticated user modeling servers [6,7]. Recent works propose a common user ontology, like GUMO [8], that creates a common syntax and understanding between all applications [9]. With the ontology as a shared data format, sharing of information is not a problem. But, this method is hardly possible in a commercial environment as most applications have application dependent user profiles. These profiles are adapted to the special requirements of each application and are firmly rooted within each system architecture. To change the structure of the user profile is always related with big changes to the system architecture, which means extra costs for the application providers.

The mediation approach [10] is a more practical approach. It aims to build a integrated user model suitable for a specific goal, e.g. recommendation of music. This integrated user model is based on data collected from different applications and aggregated by specialized software components. These software components transform data from one representation into a target representation. This approach solves the heterogeneity problem by having specialized software components for each transformation. The shortcoming here is that for each data field that should be transformed a separate component has to be developed which can lead to immense computational efforts.

3 Semantic User-Centric Data Management

In our approach, we focus on a hybrid strategy utilizing ideas from both types. To aggregate information, we developed an ontology, presented in section 3.1, as the basis for our system. For the sharing, we build a privacy-preserving framework, presented in section 3.3 and 3.4, which manages the data using different mediators and shares information between applications only with user consent. Figure 1 shows three user profiles with three attributes each, containing personal information and interest information, which should be aggregated utilizing our approach.

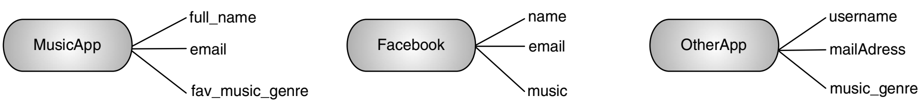


Fig. 1. Three different user profiles containing personal information and music interests

3.1 The Aggregation Ontology

We present a generic ontology, which allows us to define a degree of similarity between information and to determine the information source (the application that provided the data). Figure 2 visualizes our ontology. The ontology gives us the possibility to define a model, which describes how information in different profiles is related and how data can be aggregated. This aggregation model (AM) with descriptions about coherences between the user profile data is the basis for our system and for the in section 4 presented approach to show users their personal information stored in different applications.

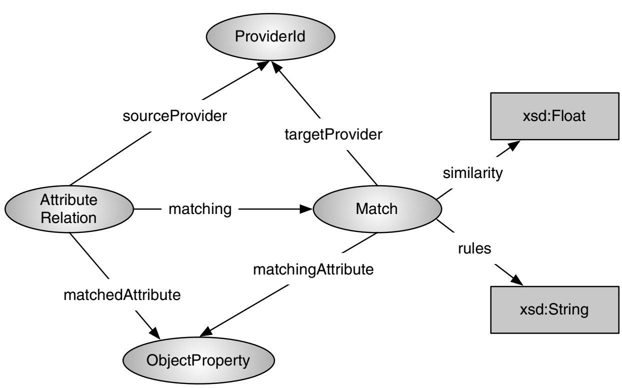


Fig. 2. Ontology for Profile Aggregation

An important extension to existing ontological approaches, e.g. [11], is the entity Match, which allows adding extra knowledge to the model. Extra knowledge can be a similarity measure of related data in two different user profiles or instructions. Instructions can be a set of predefined rules [12] describing how to aggregate information. Such extra information is needed for the aggregation of user profiles, as it is an important indication for mediators how to handle the data. Table 1 gives a detailed description all entities and relations of the ontology.

Table 1. Entities and relations of the ontology

Entity/Relation	Description
ProviderId	The description of an UM provider with name and id. E.g. Facebook.
AttributeRelation	Defines a relation between the requested attribute (matchedAttribute) of a provider (sourceProvider) with two or more attributes. Connected attributes can be from different providers or only from one provider.
Match	Encapsulates different attribute relations with extra information like similarity, or rules how to aggregate data.
owl:ObjectProperty	Defines the relations between instances of two classes.
sourceProvider	The application identifier of the application hosting the UM.
targetProvider	UM provider of the attribute similar to an requested attribute.
matching	Relation between an AttributeRelation Entity and a Match Entity.
similarity/rules	The similarity attribute defines the degree of similarity between two attributes. The rules attribute defines aggregation rules for the profiles
matchingAttribut	Defines the special application attribute that corresponds to the aggregation model attribute.
matchedAttribut	Defines the attribute that can be requested by other applications to get similar attributes from different applications.

3.2 User Profile Aggregation with the Ontology

Based on this ontology, we can define a concrete model that allows us to aggregate user profiles from different applications and access the information in a unified way. To outline the approach, we exemplarily connect the three user profiles shown in Figure 1 (MusicApp, Facebook, OtherApp) containing personal (name, mail) and interest (music) information.

The actual definition of the model is a straightforward process. First, one has to analyze the given structure of the different user profiles that should be connected. The goal is to find similar attributes in different profiles that contain similar data. For attributes where the contained information is only partly related, a similarity measure has to be defined. The similarity measure is a substantial information for the data management and visualization process as it is an important indicator for the system how to handle the data. Such a similarity definition can be done manually, semi-automatically or automatically [13]. The aggregation of the profiles can be automatedo some extent [14]. In this scenario, we perform the aggregation manually. We have two information blocks, personal information and music interests, that can be aggregated. To aggregate the music information, we define a new AttributeRelation called 'music favorite genres' in our aggregation model (AM). We define matchedAttribute (AM#music favorite genres) and sourceProvider (AM#AM ID) entries accordingly, which are needed to access the model and retrieve information.

```
<pdm:AttributeRelation rdf:ID="music_favorite_genres">
              <pdm:matchedAttribute rdf:resource="AM#music_favorite_genres"/>
              <pdm:sourceProvider rdf:resource="AM#AM_ID"/>
              <pdm:matching>
                <pdm:Match>
                  <pdm:similarity>0.9</pdm:similarity>
                  <pdm:targetProvider rdf:resource="musicApp#providerId"/>
                  <pdm:matchingAttribute
rdf:resource="musicApp#fav_music_genres"/>
                </pdm:Match>
                <pdm:Match>
                  <pdm:similarity>0.3</pdm:similarity>
                  <pdm:targetProvider rdf:resource="facebook#providerId"/>
                  <pdm:matchingAttribute rdf:resource="facebook#music"/>
                </pdm:Match>
                <pdm:Match>
                  <pdm:similarity>0.6</pdm:similarity>
                  <pdm:targetProvider rdf:resource="otherApp#providerId"/>
                  <pdm:matchingAttribute rdf:resource="otherApp#music_genres"/>
                </pdm:Match>
              </pdm:matching>
</pdm:AttributeRelation>
```

Fig. 3. Aggregation of music interests from 3 different profiles using the ontology

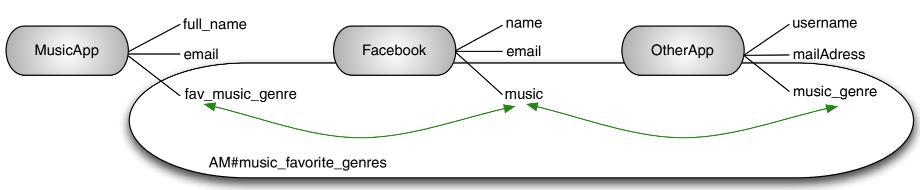


Fig. 4. The different music interests in the user profiles aggregated using the ontology

Fig. 3 shows the resulting model that describes the relations (Fig. 4) of the music interest attributes between the profiles. The aggregation of the personal information attributes follows this process. Once the model is defined, it is integrated into a framework that offers a web-service API to access the information.

3.3 Privacy Aware Data Sharing

The requirements for a privacy-aware approach sharing information safely across applications are to make sure that no personal user information is shared unintended. To fulfill these requirements, we utilize OpenID[4]. OpenID offers an interface to give permissions to third-party applications to use data and to actually share it. We have

[4] http://openid.net/

chosen OpenID because it is a well established technology supported by companies like Google and Microsoft.

3.4 Access to the Aggregated Data

The actual data access is handled by the API which offers methods to request information on behalf of the user. Such a request can come from the user, who wants to access personal data or other systems that want to use the data for personalization or adaptation purposes. All access to the data must be confirmed by the user using the OpenId interface. If a system asks for information about 'AM#music favorite genres', and the user approves the request, the system gets the information stored in the musicApp attribute 'musicApp#fav music genres', the Facebook attribute 'facebook#music' and the OtherApp data from 'OtherApp#music genres' as these are related to the 'AM#music favorite genres' field in the aggregation model.

4 My Personal User Interface

"My Personal User Interface" is a system that uses the presented ontology and framework, to visualize personal user information distributed over different applications. "My Personal User Interface" has the goal to assist users to

- keep track of applications they have,
- stay in control over their personal data,
- control the information flow of personal data.

To support people to have an overview over their applications and data the main UI is split into a top and bottom view, see Figure 5. The top view is a cover-flow element showing the different applications of a user. The cover-flow allows selecting an application and getting an overview of the personal information stored in it. This personal information is presented in the bottom view of the UI. For example, Figure 5 shows personal contact information stored in Facebook. We adopted the information card metaphor [15] to visualize the different applications of the user and to visualize the stored personal information.

Figure 6 gives an example of the type of information and how it is visualized in our system. The user has different information cards visualizing information stored about her. The user can see her last actions in an application, as an example of implicit information visualized by our system.

We also designed an UI to support users in privacy-aware use-cases, see Figure 7, which allows users to control what kind of data is distributed. If a third-party application requests data of the user, the user is asked to give permission for that application to use the data. Therefore, the UI presents data that will be sent to the user. The process of sharing information is two-folded. After a third-party application requests data, e.g. about the user's musical taste, the system selects all information previously aggregated using the ontology. So, for a music taste request, data from Facebook or the previously described 'Profile Analysis'-profile would be selected. The second step is the validation through the user. Therefore, the selected information is presented to the user using the same card metaphor but only showing the information to be sent to the

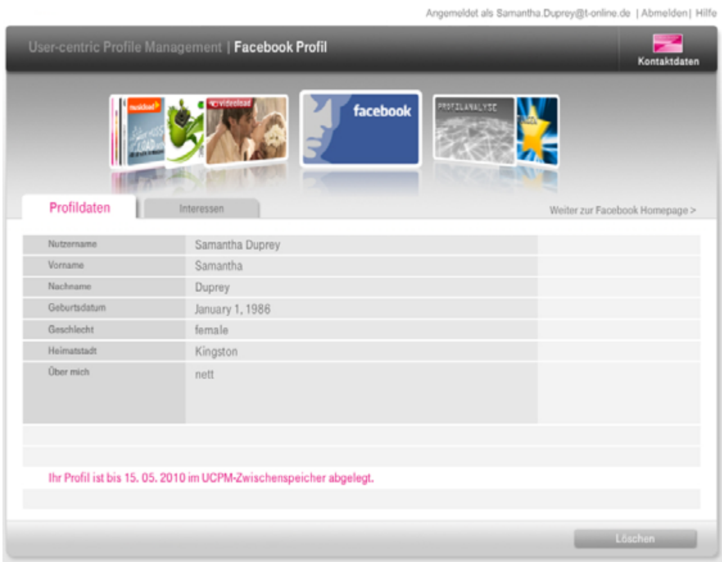


Fig. 5. Main UI showing different applications and personal information

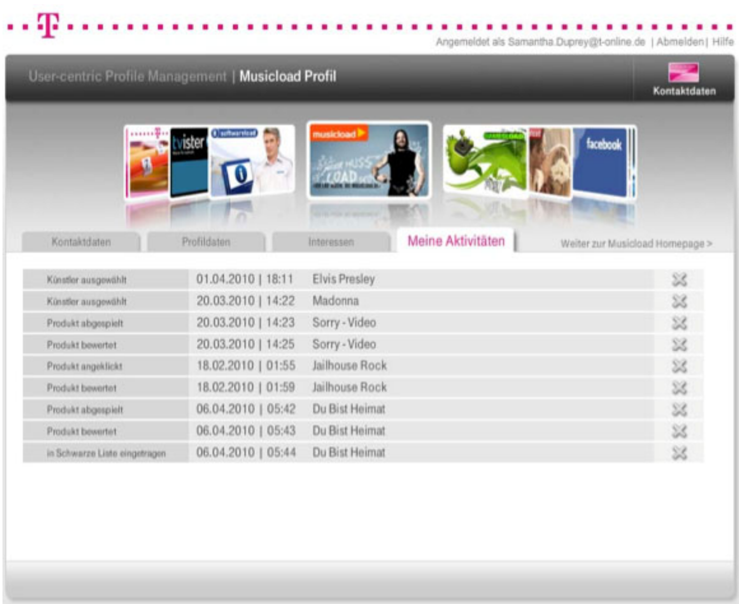


Fig. 6. Visualization of past behavior

third-party application. The user can navigate through the different cards, see which information will be sent and decide to accept or deny the request. This makes it possible to easily see and control what data is shared.

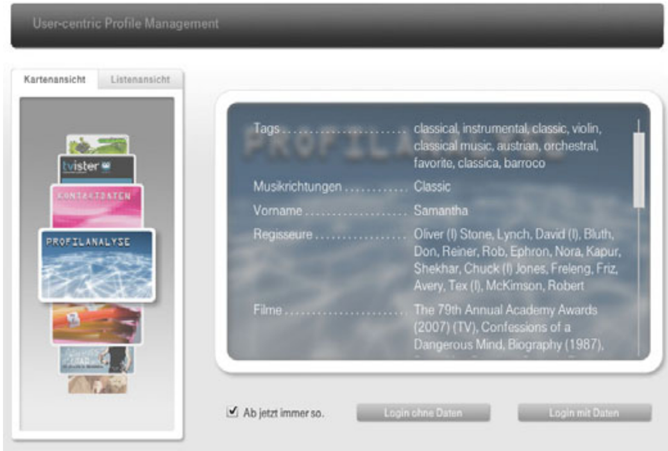


Fig. 7. Our OpenId interface. Buttons at the bottom allow users to accept or deny a data request.

5 Conclusion and Outlook

In this paper we introduced a solution to aggregate, access and manage personal information in a secure way. We presented a new ontology that supports the aggregation of distributed user profiles and a framework that utilizes the ontology and allows secure data sharing by using OpenId. On top of this framework, we created My Personal User Interface, a system that allows users to fully profit from the semantic framework and to keep an overlook over their personal data. In future steps, we want to to use the aggregated user information to further enrich the knowledge about the user and add new information retrieved from the Semantic Web. This information can be used to improve personalization and adaption and help users to profit more from such a personalized system. Of course, the user stays in control and can decide to discard this automatically added information.

References

1. Abel, F., Henze, N., Herder, E., Krause, D.: Interweaving Public User Profiles on the Web. In: De Bra, P., Kobsa, A., Chin, D.N. (eds.) UMAP 2010. LNCS, vol. 6075, pp. 16–27. Springer, Heidelberg (2010)
2. Plumbaum, T.: Semantically-Enhanced Ubiquitous User Modeling. In: De Bra, P., Kobsa, A., Chin, D.N. (eds.) UMAP 2010. LNCS, vol. 6075, pp. 407–410. Springer, Heidelberg (2010)
3. Plumbaum, T., Lommatzsch, A., Rudnitzki, S., Luca, E.W.D., Dwiger, H., Albayrak, S.: Adaptive music news recommendations based on large semantic datasets. In: 1st Workshop On Music Recommendation And Discovery (WOMRAD), ACM RecSys, 2010, Barcelona, Spain (2010)

4. Kuflik, T.: Semantically-enhanced user models mediation: Research agenda. In: 5th International Workshop on Ubiquitous User Modeling (UbiqUM 2008) (January 2008)
5. Finin, T.W.: Gums: A general user modeling shell. In: Kobsa, A., Wahlster, W. (eds.) User Models in Dialog Systems, pp. 411–430. Springer, Heidelberg (1989)
6. Kay, J., Kummerfeld, B., Lauder, P.: Personis: A Server for User Models. In: De Bra, P., Brusilovsky, P., Conejo, R. (eds.) AH 2002. LNCS, vol. 2347, pp. 203–212. Springer, Heidelberg (2002)
7. Fink, J.: User Modeling Servers Requirements, Design, and Evaluation. PhD thesis, Universitt Duisburg-Essen (2003)
8. Heckmann, D., Schwarzkopf, E., Mori, J., Dengler, D., Krner, A.: The user model and context ontology gumo revisited for future web 2.0 extensions. In: Bouquet, P., Euzenat, J., Ghidini, C., McGuinness, D.L., Serafini, L., Shvaiko, P., Wache, H. (eds.) Proceedings of the International Workshop on Contexts and Ontologies: Representation and Reasoning (CO:RR). CEUR Workshop Proceedings, CEUR-WS.org, vol. 298 (2007)
9. Gruber, T.R.: Toward principles for the design of ontologies used for knowledge sharing. Int. J. Hum.-Comput. Stud. 43(5-6), 907–928 (1995)
10. Berkovsky, S., Kuflik, T., Ricci, F.: Mediation of user models for enhanced personalization in recommender systems. User Modeling and User-Adapted Interaction 18(3), 245–286 (2008)
11. van der Sluijs, K., Houben, G.J.: A generic component for exchanging user models between web-based systems. Int. J. Continuing Education and Liflong Learning 16(1/2), 64–76 (2006)
12. Leonardi, E., Abel, F., Heckmann, D., Herder, E., Hidders, J., Houben, G.J.: A Flexible Rule-Based Method for Interlinking, Integrating, and Enriching User Data. In: Benatallah, B., Casati, F., Kappel, G., Rossi, G. (eds.) ICWE 2010. LNCS, vol. 6189, pp. 322–336. Springer, Heidelberg (2010)
13. Banerjee, S., Pedersen, T.: Extended gloss overlaps as a measure of semantic relatedness. In: Proceedings of the 18th International Joint Conference on Artificial Intelligence, pp. 805–810. Morgan Kaufmann Publishers Inc., San Francisco (2003)
14. Noy, N.F.: Semantic integration: a survey of ontology-based approaches. SIGMOD Rec. 33, 65–70 (2004)
15. Ballinger, K., Bill Barnes, G.S., Causey, J.: Patterns for supporting information cards at web sites: Personal cards for sign-up and sign-in. Technical report, Microsoft (2007)

Part VII

Measuring and Recognising Human Behaviour

Effect of Menstrual Distress on Task Performance

Keiko Kasamatsu[1], Mi Kyong Park[2], and Seiko Taki[2]

[1] Division of Industrial Art, Faculty of System Design, Tokyo Metropolitan University
[2] Division of Management Systems Engineering, Faculty of System Design, Tokyo Metropolitan University, 6-6, Asahigaoka, Hino, Tokyo, Japan
{kasamatu,mkpark}@sd.tmu.ac.jp, staki@cc.tmit.ac.jp

Abstract. Women have periodic physiologic change of menstruation. There is a possibility to be influenced the physical and mentally in the female. It is necessary to support a comfortable life and work for female. The relationship between workload and stress that was taken the menstrual cycle into consideration had been examined on mental tasks in this research. Menstrual phase, follicular phase and premenstrual phase of menstrual cycle were focused and analyzed. The survey was conducted to grasp the menstrual cycle of the subjects. The experiment was consisted of two kinds of tasks. As the results of the survey and the experiment, it was suggested that there were differences on task speed, stress, and sympathetic and parasympathetic nerve activity due to menstrual complaints.

Keywords: Menstrual cycle, Task Performance, Mental Workload, Menstrual Associated Symptoms.

1 Introduction

There are many opportunities using information and communications technology like personal computer in work environment in late years. Information terminal uses on every workplace not only office work. The work with mental workload increases and mental stress becomes large. On the other hand, the equality on labor had been ensured by the Equal Employment Opportunity Law in Japan. As the employment situation of Japan, the female worker increases and the scope of female occupations has expanded. The equality on labor is very important, however it is thought that there are problems when a biologically different man and woman does physical work.

Women have periodic physiologic change of menstruation. There is a possibility to be influenced the physical and mentally in the female. Similar physical work to the man will force the large encumbrance for such a woman. Previous researches indicated that some women suffer from Premenstrual Syndrome (PMS) and Premenstrual Dysphoric Disorder (PMDD)[1]. Grady-Weliky[2] reported that 75 percent of women of reproductive age had the influence of menstrual cycle like premenstrual symptoms at some time during their lives. Therefore, it is important issue on occupational health

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 597–602, 2011.

and security to clarify the characteristic labor stress of female. It is necessary to support a comfortable life and work for female.

The relationship between workload and stress that was taken the menstrual cycle into consideration had been examined on mental tasks in this research.

2 Methods

Menstrual phase, follicular phase and premenstrual phase of menstrual cycle were focused and analyzed. The survey was conducted to check the menstrual cycle of the subjects. The experiment was consisted of two kinds of tasks.

2.1 Survey

The surveillance period was from the start of menstruation to the end of next menstruation. The body basal temperature was taken and recorded every morning for the surveillance period and the mental and physical complaints were investigated by the questionnaire every day. The subjects were fifteen, female college students.

The MDQ (Menstrual Distress Questionnaire) was developed to create a questionnaire about menstrual cycle complaints by Moos[3] and the questionnaire used in this research was modified the MDQ. The modified MDQ with 54 items in Japanese was administered to investigate physical and mental conditions every day in this survey.

2.2 Experimental Methods

As a result of the survey, it experimented for ten subjects who were the menstrual cycle within normal range days. The three phases for experiment were menstrual phase, follicular phase and premenstrual phase. Three times every subject were conducted the experiment one time for each phase. The subject continued to record the body basal temperature and fill out the questionnaire.

2.3 Experimental Tasks

The two kinds of experimental tasks were one digit input task (task A) and a mental arithmetic task involving the addition of two digits (task B). The experimental time was fifteen minutes for each task. Task A was input task which inputting the numeral of one digit randomly displayed for the screen from the ten-key. Next numeral is not displayed, until the right numeral is input. Task B was mental arithmetic task involving the addition of two digits. The addition problem was shown on screen, and there was no answer that the same figure queues up such as 22, 33. The screen did not advance to the next problem until a correct answer was entered using the number pad. If the answer was incorrect, the system waited for the correct answer to be entered.

2.4 Measurement indices

The measurement indices were time required to answer one question (response time) and error numbers. The modified MDQ, physical condition investigation and stress measurement by salivary amylase were conducted on before and after experiment. In addition, electrocardiograph (ECG) was measured to examine the mental workload during experiment. ECG data measured for two minutes before each experiment and for fifteen minutes during the experiment. The frequency analysis was done to these data, and LF/HF was obtained.

3 Results

3.1 Results on the Survey

The respondent to a survey were fifteen. The analytical subjects were ten with normal menstrual cycle.

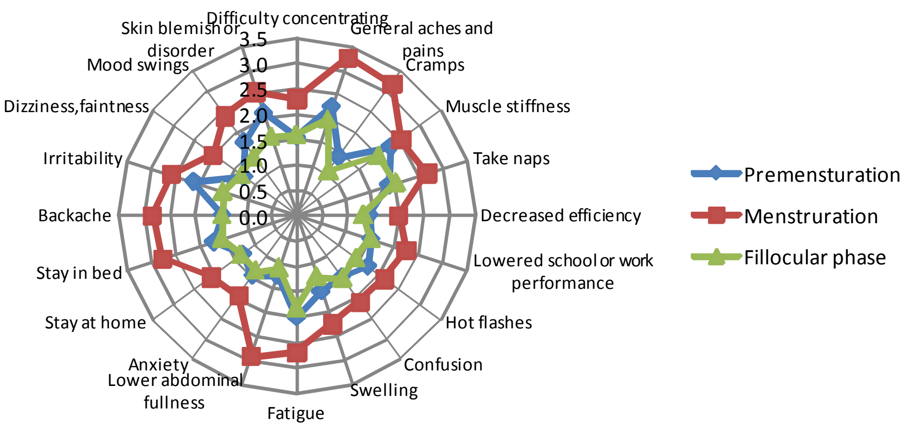


Fig. 1. The score on 21 items of MDQ

The average of menstrual cycle days was 28.8 days. The average of menstruation days were 5.6 days. The average of basal body temperature was 36.40 degrees. The average of sleeping hours was 6.75 hours. The average of drinking days at a menstrual cycle was 0.12 days. The score of modified MDQ on menstrual phase was the highest. The score of 21 items was high on menstrual phase. 21 items were difficulty concentrating, general aches and pains, cramps, muscle stiffness, take naps, confusion, and so on as Fig.1. There were the symptoms as irritability and skin blemish or disorder in premenstrual phase. It was suggested that it was necessary to manage about health enough especially, and to take care about the stress similar to general view. However, the management matched to an individual symptom and the extent is necessary so that there is an individual variation in the menstrual cycle and the menstruation accompaniment symptom.

Table 1. With and without the symptoms in eight factors on premenstrual and menstrual phase

		Menstruation			
		with symptoms		no symptoms	
Premenstruation	with symptoms	Pain	5		
		Negative affect	6	Negative affect	2
		Concentration	2	Concentration	1
		Behavioral change	1	Behavioral change	1
		Autonomic reactions	2		
		Water retention	5		
		Arousal	1	Arousal	1
	no symptoms	Pain	4	Pain	1
				Negative affect	2
		Concentration	2	Concentration	5
		Behavioral change	5	Behavioral change	3
		Autonomic reactions	4	Autonomic reactions	4
		Water retention	2	Water retention	3
		Arousal	3	Arousal	5

MDQ is classified into eight factors. The eight factors were Pain, Negative affect, Concentration, Behavioral change, Autonomic reactions, Water retention, Arousal. In this research, the score of symptom item was calculated on premenstrual and menstrual phase. The subjects were classified by with or without symptoms on premenstrual and menstrual phase. The case with symptoms was the higher score than follicular phase significantly. As the results(Table 1), there were symptoms of Pain in 50% of subjects on both premenstrual and menstrual phase. 40% of the subjects had pain symptoms on menstrual phase and not on premenstrual phase. There were many symptoms of Negative affect on both premenstrual and menstrual phase. There was a little symptom of Concentration on both phases. There was no symptom of any factor on both premenstrual and menstrual phase for subject J.

3.2 Results on the Experiment

The differences of degrees on symptoms were examined on eight factors of MDQ and each phase.

For response time, there was no significant difference on eight factors and phase of task A. The statistical significant difference was not recognized on task B. However, the response time on group with symptoms on premenstrual and menstrual phase was about 1000ms longer than on group without symptoms on both phases(Figure 2,3). In particular, the response time was influenced by the factors of negative affect, concentration, autonomic reactions, and water retention.

The error rate was compared on factors and phases. As the results, there was no significant difference on two kinds of tasks.

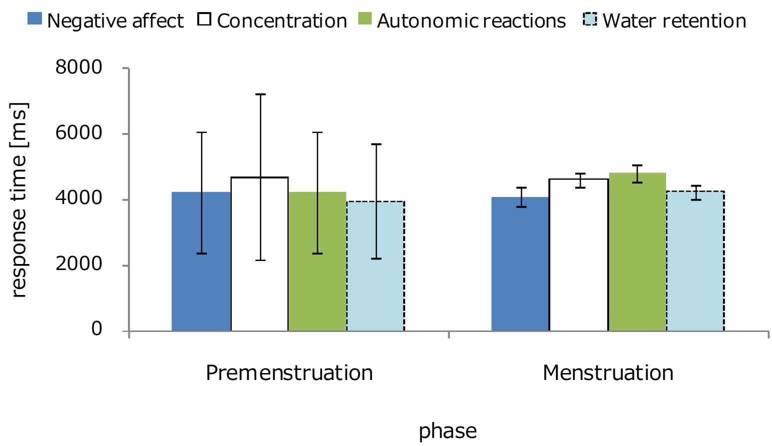


Fig. 2. The response time on group with symptoms on premenstrual and menstrual phase

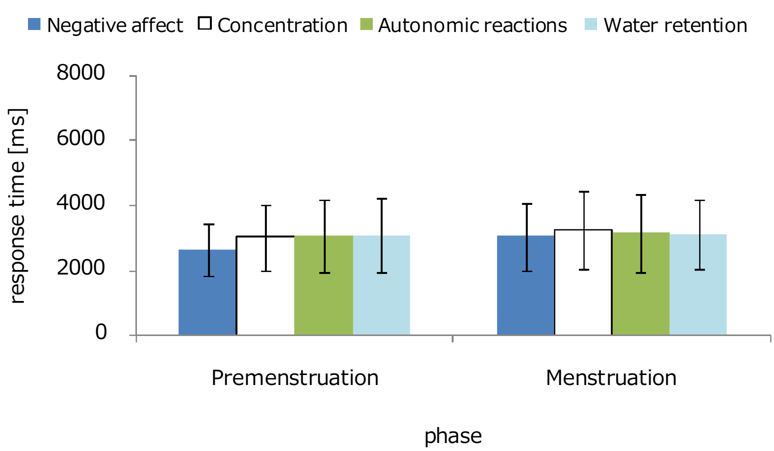


Fig. 3. The response time on group without symptoms on premenstrual and menstrual phase

There were significant differences on some factors and phases for stress test by salivary amylase. The results of a group which symptoms appeared strongly by premenstrual and menstrual phase was as follows; there were significant differences between premenstrual phase and follicular phase($p<0.05$) on Pain, and between premenstrual phase and follicular phase($p<0.05$) on Water retention for Task A. Moreover, there was a significant difference between premenstrual phase and menstrual phase($p<0.05$) on autonomic reactions for Task B.

LF/HF component was calculated from electrocardiographic data. The rate of increase on LF/HF component on Task B of premenstrual phase was higher than Task A. The rate of increase on it on Task A of menstrual phase and follicular phase was higher than Task B.

4 Discussion

There were many menstrual associated symptoms on menstrual phase. There was individual difference on the degree of symptoms.

For the task performance, the response time on group with symptoms of premenstrual and menstrual phase was longer than on group without symptoms of both phases on task with thinking. There was hardly an influence due to menstrual associated symptoms on simple task like Task A. It was suggested that there was an influence for task speed due to menstrual associated symptoms on thinking task.

For stress test, it was recognized that the stress on menstrual phase was large on Task A. As for the results of LF/HF component, stress condition on the Task B of premenstrual cycle was more stronger than Task A, and it on the Task A of menstrual phase and follicular phase was more stronger than Task B.

5 Conclusions

This research was examined the differences between phases on the extent of menstrual associated symptoms, task performance, and physiological indices by grouped with or without menstrual complaints. As the results, it was suggested that there were differences on task speed, stress, and sympathetic and parasympathetic nerve activity due to menstrual complaints.

References

[1] Dalton, K., Holton, W.: Once a month -Understanding and treating PMS. Hunter House Inc., Alameda (1999)

[2] Grady-Weliky, T.: Premenstrual dysphoric disorder. N. Eng. J. Med. 348(5), 433–438 (2003)

[3] Moos, R.H.: The Development of a Menstrual Distress Questionnaire. Psychosom. Med. 30(6), 853–867 (1968)

A Study of Attention Control by Using Eye Communication with an Anthropomorphic Agent

Tatsuya Mita, Ryo Wada, Noriaki Kuwahara, and Kazunari Morimoto

Graduate School of Science and Technology, Kyoto Institute of Technology, Matsugasaki, Sakyo-ku, Kyoto, Japan
tatsuya_tatsuya_mi@yahoo.co.jp, ryo.wada@gmail.com, {nkuwahar,morix}@kit.ac.jp

Abstract. We are developing the self-toileting support system for the dementia patient by combining posture-detection method and video contents. We conducted a preliminary evaluation, and found a problem. When we gave instruction to pick up the paper in the box to patient by video-voice guidance, the patient persisted in the box in the audio-visual guidance and did not notice the real box. From this result, we noticed that it was difficult for people with dementia to shift the eye gaze and attention from something to something only by giving verbal instruction. We are investigating the way of the attention control by using the eye communication model with the anthropomorphic agent. In this paper, we showed the results of the experiment in that we compared the effectiveness of the eye gaze and the attention control by adding the text message, the arrow animation, and the agent animation respectively to the voice instruction.

Keywords: Dementia, Anthropomorphic agent, Eye communication model, Attention shift, Joint attention, Eye gaze.

1 Introduction

The number of people with dementia is increasing rapidly all over the world. Assistances of toileting, meal and bathing are known as the three major burdened cares. Among those, toileting assistance is considered the most stressful for caregivers. Furthermore, for the patient at home who wants self-toileting, toilet failure is not only the hygiene issue, but also hurts his/her self-esteem. We are developing the self-toileting support system for the dementia patient by combining posture-detection method and video contents [1]. The posture-detection system using IR cameras and invisible IR dot pattern projector detects the user's state according to his/her 3-D appearance (Fig. 1). Video and voice instructions consist of scenes describing nine-step actions, such as asking the patient to stand in front of the toilet bowl, and then pull down his pants, and so on.

We conducted a preliminary experiment, and found a problem. When we gave him/her an instruction to pick up the paper in the box by video-voice guidance, the patient focused on only the box in the audio-visual guidance and did not notice the

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 603–611, 2011.

real box. From this result, we noticed that it was difficult for people with dementia to shift their eye gaze and change attention only by giving verbal instruction.

Nakamoto et al. conducted the experiment where they compared the effect of the eye ball animation with that of the arrow animation for inducing the eye gaze shift [2]. They showed that the eye ball animation was better than the arrow animation for the elderly people in order to shift their eye gaze because the eye ball animation induced the reflective eye movement.

Yonezawa et al. proposed the eye communication model between a human and an anthropomorphic agent. In the eye communication model, a joint attention is followed by an eye contact between the human and the agent. They showed that the eye communication model improved the reliability and the acceptance of the information presented by the agent [3].

These researches indicate the possibility of developing the effective way for the eye gaze and attention control of a person with dementia.

In order to overcome the attention control problem of a person with dementia, we have been investigating the way of the attention control by using the eye communication model with the anthropomorphic agent [4]. In this paper, we showed the results of the experiment in that we compared the effectiveness of the eye gaze and the attention control by adding the text message, the arrow animation, and the agent animation respectively to the voice instruction.

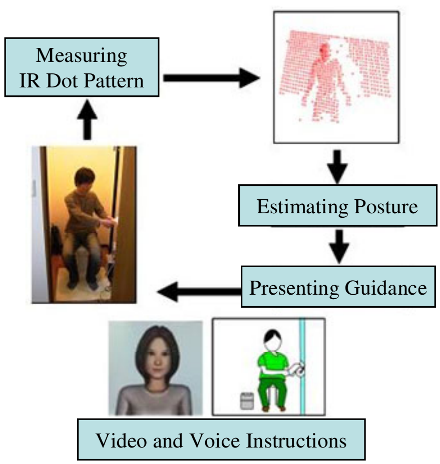


Fig. 1. Overview of Self-toileting support system

2 Experiment

2.1 Outline

The purpose of this experiment was to examine the effectiveness of the eye gaze and the attention control by using the reflective eye movement and the eye communication model. The baseline was the voice instruction that asked the subject to push the left or right button. We added the text message, arrow animation, and the agent animation respectively to the voice instruction. The agent animation was designed to perform the eye communication model and to induce the reflective eye movement. We presented these stimuli to the subject randomly, asked the questionnaires, and measured the response time to push button.

2.2 Method

Fig.2 shows the experimental environment. We set up the display and the speaker for presenting the instructions in front of the subject, and two buttons for responding to the instructions on the left and right of the subject. Two video cameras were installed in order to record the behavior of the subjects.

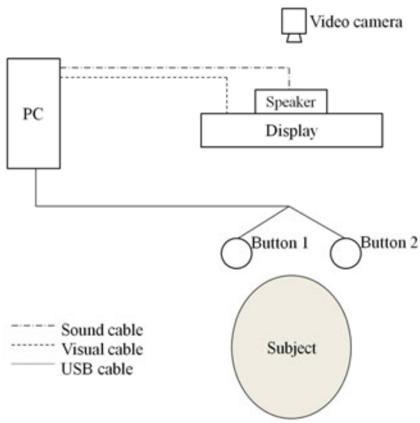


Fig. 2. Experimental environment

Fig. 3 shows the visual instructions used in the experiment. In each instruction, we asked the subjects to push the left button or the right button by the voice. Fig. 4 shows an example of visual stimulus of the arrow animation. Firstly, the small circle was flashed three times per second after the blank screen was displayed for one second. Then, the visual instruction was displayed with the voice instruction at the same time. We also examined the case where no visual stimulus was presented with the voice instruction as the baseline. In this case, the blank screen was displayed while the voice instruction was notified. After the instruction the screen was hold for three seconds, and it returned to the blank screen. Then, the next instruction was presented in the same manner as mentioned above.

Before the experiment, we explained the experimental protocol to the subjects: they were asked to push the right button or the left button according to the instruction. One session consisted of ten instructions included randomly the left instruction and the right instruction, and total 4 sessions were performed by each subject. The subjects answered a questionnaire to evaluate friendly, reliable, comfort, and so on subjectively in every session. During the experiment, the subjects sat on a chair, and the instruction was presented automatically to the display.

Subject. The subjects were healthy young and elderly people. The young subjects were 20 students of KIT University. The age ranged from 22 years to 30 years (10 male, 10 female). The elderly subjects were 20 people who live in Kyoto city. The age ranged from 63 years to 76 years (10 male, 10 female).

Evaluation Form. Table 1 shows the questionnaire items. We selected ten adjective terms, and each term was evaluated by five-grade evaluations ("none feel at all" - "no feel" - "neither" - "yes" - "extremely yes"). In addition, after the experiment, we asked the subject to rank the types of the instructions: only the voice, the voice with the text, the voice with the arrow animation, the voice with the agent animation.

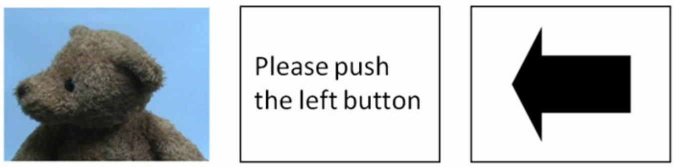


Fig. 3. The three types of visual instructions used in the experiment (from left to right): anthropomorphic agent, text instruction, arrow

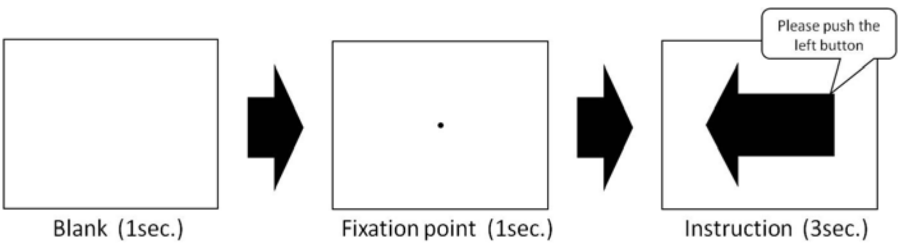


Fig. 4. An example of the visual stimulus of the arrow animation

Table 1. Questionnaire items

Number	Item
Q1	Friendly
Q2	Reliable
Q3	Uncomfortable
Q4	Tired
Q5	Easy
Q6	Intuitive
Q7	Unnatural
Q8	Assured
Q9	Safe
Q10	Unconscious

2.3 Results

Fig. 5 shows the questionnaire results in the case of young people, and Fig. 6 shows the questionnaire results of elderly people. Each figures show an average (μ) and a standard deviation (σ) of each questionnaire item of each instruction. In addition, outliers (outside the μ±2σ) have been removed.

We performed a one-way ANOVA (repeated-measures design) for each questionnaire item. The statistically significant portions are shown Fig. 5 and Fig. 6 ($p<.05$). As for the young people, the anthropomorphic agent was higher in seven items: "Friendly", "Reliable", "Tired", "Easy", "Intuitive", "Safe", and "Unconscious". In the case of elderly people, the anthropomorphic agent was higher in two items of "Friendly" and "Reliable".

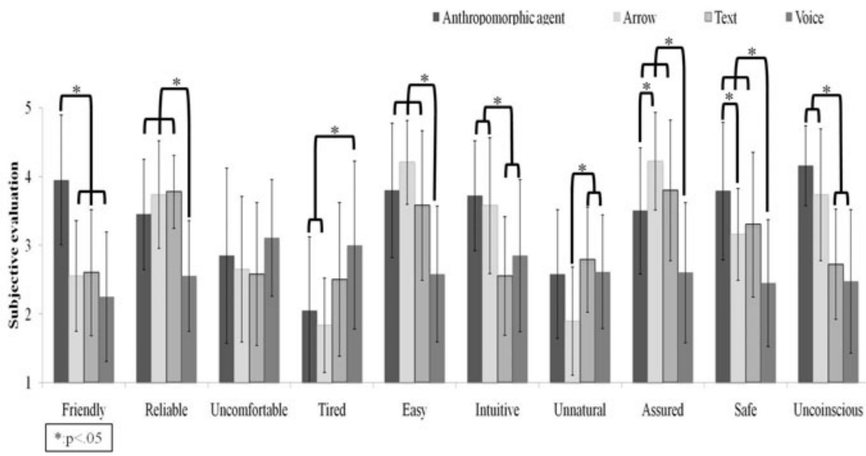


Fig. 5. The questionnaire results in the case of young people

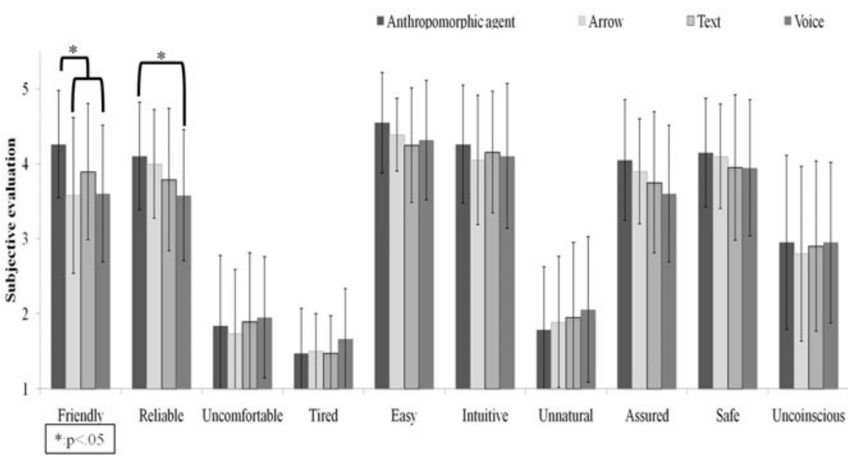


Fig. 6. The questionnaire results in the case of elderly people

2.4 Factor Analyses

We also performed a factor analyses to investigate the factors hidden in the questionnaire results. The factors of each instruction were extracted in the case of both young people and elderly people. Table 2 shows the analysis result of the anthropomorphic agent and the arrow instructions of elderly people. Also, Table 3 shows those of young people.

As for the young people on the instruction of the anthropomorphic agent, they focused on these items by factor 1: "Assured", "Easy" and "Reliable". On the other hand, as for the elderly people on the instruction of the anthropomorphic agent, they focused on these items by factor 1: "Unconscious", "Assured", "Reliable", "Safe" and "Friendly".

Table 2a. Factor analysis results of questionnaire investigation in the case of young people (factor loading after varimax rotation) (a) Anthropomorphic agent

	I	II	III	Communality
Assured	**.898**	.013	-.022	.806
Easy	**.782**	-.050	-.136	.633
Reliable	**.702**	.262	-.063	.565
Friendly	-.108	**.980**	-.129	.989
Safe	.303	**.793**	-.089	.728
Tired	-.182	**.569**	.316	.456
Unconscious	.198	.341	-.044	.157
Unnatural	-.209	-.102	**.882**	.832
Uncomfortable	-.191	-.162	**.777**	.666
Intuitive	.149	.140	.309	.137
Factor contribution	2.188	2.156	1.625	5.969
Cumulative contribution ratio	21.88	43.45	59.69	

Table 2b. Factor analysis results of questionnaire investigation in the case of young people (factor loading after varimax rotation) (b) Arrow

	I	II	III	Communality
Unconscious	**.730**	.337	.026	.648
Assured	**.654**	.267	-.392	.653
Reliable	**.645**	.390	.114	.581
Safe	**.622**	-.131	-.072	.409
Uncomfortable	-.581	-.235	.126	.409
Intuitive	.203	**.868**	-.089	.802
Easy	.230	**.609**	-.215	.470
Unnatural	-.312	-.286	**.752**	.744
Tired	.148	-.073	**.700**	.517
Friendly	**.438**	-.483	-.579	.761
Factor contribution	2.507	1.854	1.633	5.995
Cumulative contribution ratio	25.07	43.61	59.95	

Table 3a. Factor analysis results of questionnaire investigation in the case of elderly people (factor loading after varimax rotation) (a) Anthropomorphic agent

	I	II	III	Communality
Unnatural	**.910**	-.085	-.356	.963
Tired	**.818**	-.362	-.341	.916
Uncomfortable	**.754**	-.199	-.083	.615
Friendly	-.120	**.861**	.269	.828
Intuitive	-.300	**.784**	.175	.736
Easy	-.210	**.657**	.394	.631
Assured	-.260	.240	**.931**	.991
Reliable	-.238	**.490**	**.682**	.761
Safe	-.535	**.440**	**.547**	.780
Unconscious	.081	-.086	-.269	.086
Factor contribution	2.632	2.464	2.211	7.307
Cumulative contribution ratio	26.32	50.96	73.07	

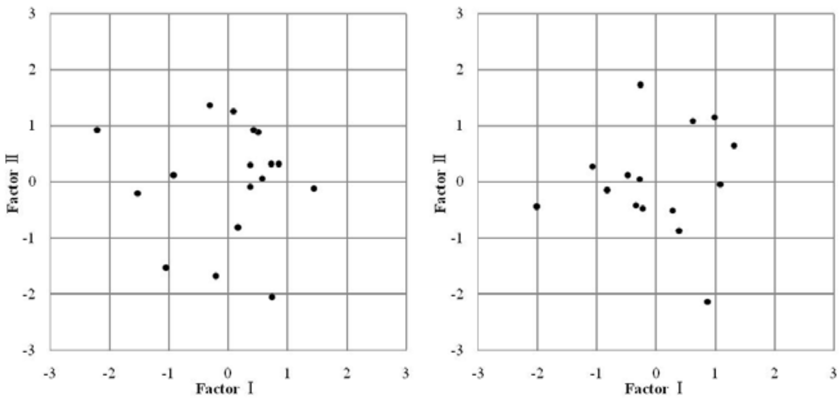
Table 3b. Factor analysis results of questionnaire investigation in the case of elderly people (factor loading after varimax rotation) (b) Arrow)

	I	II	III	Communality
Reliable	**.913**	.083	.283	.921
Unnatural	-.783	**.414**	.063	.789
Assured	**.693**	-.376	.148	.643
Friendly	**.591**	-.400	**.498**	.758
Safe	**.555**	-.361	**.431**	.624
Tired	.005	**.813**	-.132	.678
Uncomfortable	-.435	**.786**	-.055	.809
Unconscious	-.173	.344	-.126	.164
Easy	-.030	-.225	**.816**	.717
Intuitive	.370	.005	**.782**	.748
Factor contribution	2.941	2.057	1.853	6.851
Cumulative contribution ratio	29.41	49.98	68.51	

3 Discussion

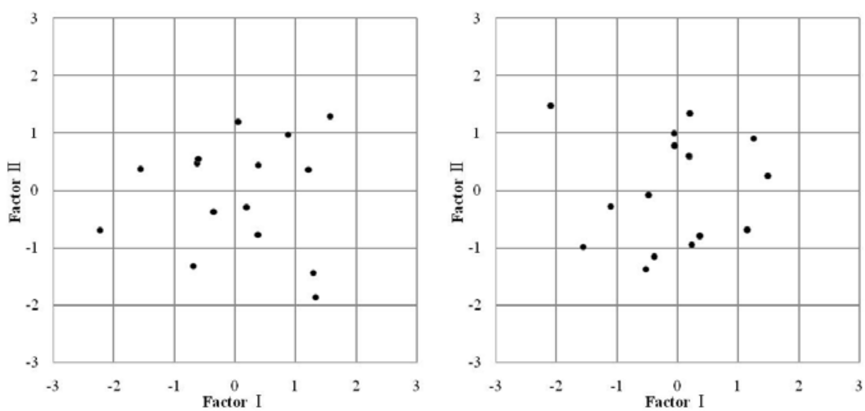
As for the young people, there is a significant difference between the voice only instruction and the other instructions in four items. The subjects felt worse to the voice only instruction. There is another significant difference between the instruction of anthropomorphic agent and the other types of instruction in "Friendly", "Safe", and "Unconscious". The presence of anthropomorphic agent gave such feelings to the subject more than the other type of the instruction. On the other hand, the arrow instruction was evaluated as the best in "Assured". Therefore, the anthropomorphic agent created good atmosphere for the subject, but in order to understand the instruction surely, the arrow instruction was the best.

The three factors for the agent instruction were obtained as the results of a factor analysis as shown in Table 2. Fig. 7(a) shows the scatter plot of the standardized factor scores of the agent instruction of each subject in the factor 1 and factor 2 spaces. This figure shows that the scores of all subjects are distributed with uniformly. Fig. 7(b) shows the scatter plot of the arrow instruction of the factor scores in the factor 1 and factor 2 spaces. We can see from this figure that the scores of the majority of the subjects are around zero in term of factor 2. It indicates that the subjects had the same feeling to the arrow animation in terms of easy-understanding of the instruction about the arrow instruction.



(a) Anthropomorphic agent (b) Arrow

Fig. 7. The distribution of the standardized factor scores of the young people in the factor 1 and 2



(a) Anthropomorphic agent (b) Arrow

Fig. 8. The distribution of the standardized factor scores of the elderly people in the factor 1 and 2

As for elderly people, there is the significant difference between the agent instruction and the voice only instruction on the items of "Friendly" and "Reliable". However, there are no significant differences in other items.

The three factors of the agent instruction were obtained as the results of a factor analysis shown in Table 3. Figure 8(a) shows the scatter plot of the agent instruction of the factor scores in the factor 1 and factor 2 spaces. It shows that that the scores of all subjects were distributed with uniformity. Figure 8(b) shows the scatter plot of the agent instruction of the factor scores in the factor 1 and factor 2 spaces. There were a few subjects who scored in the second quadrant, and it indicates that the almost all subjects felt good atmosphere to the arrow instruction.

4 Conclusion

In this paper, we reported the experimental results on the attention control by using the eye communication with the anthropomorphic agent. According to the questionnaire result, the young subjects preferred to the voice instructions with the agent animation. The elderly subjects showed the same tendency, but the difference between the types of the instructions were not significant in many items of the questionnaire. The factor analysis result showed that the young subjects focused on explicitness of the agent instruction, and the elderly subjects focused on familiarity of those. This indicates that the design of the appearance of the agent is more important for the elderly people.

For future works, we would like to apply to people with dementia the eye communication model with the agent, and would like to utilize this communication model in our daily living assistance system for people with dementia.

Acknowledgments. This research was supported by a grant-in-aid for scientific research (B) (21300043, 2009-2011) from Japan Society for the Promotion of Science (JSPS).

References

1. Yasuda, K., Okazaki, Y., Utsumi, A., Yamazoe, H., Abe, S.: Toilet task support system using 3-D human posture measurements: System evaluation in simulation environment. JSAI, 313–316 (2008) (in Japanese)
2. Nakamoto, S., Kato, M., Asama, H., Otake, M.: The Effect of Aging on Gaze Recognition. JSAI, pp. 313–319 (2008) (in Japanese)
3. Yonezawa, T., Yamazoe, H., Utsumi, A., Abe, S.: Evaluation of gaze-communicative stuffed-toy robot with joint attention and eye-contact, IEICE Technical Report HCS2007-48 (2007) (in Japanese)
4. Yasuda, K., Imai, Y., Kuwahara, N., Morimoto, K.: Transfer for eye gaze from screen to outer space: for dementia's toilet task support system. JSAI (2009) (in Japanese)

Auditory Feature Parameters for Music Based on Human Auditory Processes

Masashi Murakami and Toshikazu Kato

Chuo University,
1-13-27, Kasuga, Bunkyo-ku, Tokyo, Japan
{masa_m,kato}@indsys.chuo-u.ac.jp

Abstract. Authors aim to show the similarity and difference of sensibility effect when human listen to the music. For concrete example, we assume the music recommendation services. To propose the music that users really want, we use the image words that users feel for the music, and retrieve and propose the music based on the similarity of the image word. We intend to propose the music that users really want based on Kansei engineering. We make the Hierarchical model of Kansei. This model shows that human have the 4 processes when listen the music. Feature parameters are physical frequency data with physiological process. These parameters make us possible to represent how human listen the music. In next step, we represent how human feel for the music. These representations help users possible to retrieve the music really want. For representation of human feelings, we use image words and subjective estimation.

Keywords: Music, Hierarchical model of Kansei, Auditory characteristic.

1 Introduction

When human want to listen to the music, they have some needs of image. For example, when someone tired for working retrieve the music, he wants to freshen up, or concentrate. When someone dinking tea retrieve the music, he want the music that he feel "fresh" image for. We define the music that users really want is that the users feel image for. To propose the music that users really want, we use the image words that users feel for the music, and retrieve and propose the music based on the similarity of the image word. To realize this method, we make computers to trace how human listen the music, and how human feel for the music. Recently, the techniques intended for music retrieval mainly use the text-based metadata information that title, composers, and genres, as retrieval key. These services require a large amount of cost for giving metadata. Actually, quite a few systems are classified according to music itself. Therefore users retrieve the music by not the music itself but the music's metadata information. The other techniques use the frequency-based data. In these techniques, users can retrieve the music by the similarity that presumed from the frequency of the music. For presuming the similarity of the music, various algorithms are used without how human feel for the music. Therefore the evaluation of these

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 612–617, 2011.

techniques only intends to clarify the similarity of the frequency in music. These techniques are difficult to retrieve music users really want.

We intend to propose the music that users really want based on Kansei engineering. In Kansei engineering, we notice that the individualities of the basis of interpretation appear in the process of human perception and conduct a further study of the engineering modeling of the process of human perception. We make the Hierarchical model of Kansei. This model shows that human have the 4 processes when listen the music. This model, auditory processing process model shows human listen the music. In this model, we perceive the physical signals for physiological stimulus, and output the information weighted by similarity of images.

2 Hierarchical Model of Kansei

We notice the individual difference of the standard of interpretation which appears through the process of human perception and we are seeking to conduct a further study of the engineering modeling of this phenomenon (1) to (5). By the application of this way of thinking, a model of human beings and the sound is shown hierarchically (Fig.1.). This Model is divided in several levels. There are physical level, physiological level, psychological level, and cognitive level.

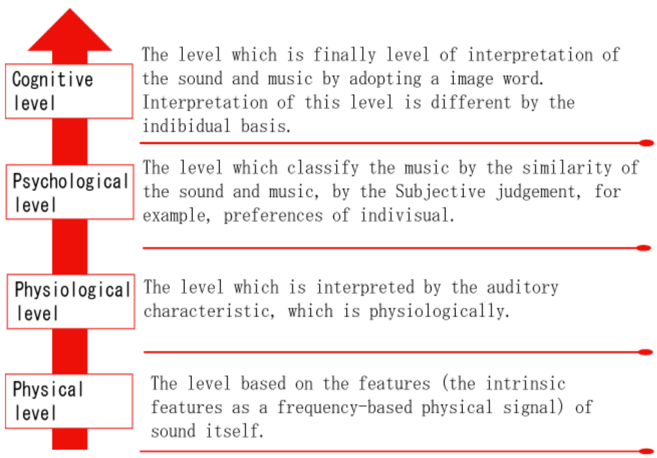


Fig. 1. Hierarchical model of Kansei for Auditory Information Processing

At first, we can find physical level. This level is based on the features (the intrinsic features as a frequency-based physical signal) of sound itself. This level deals in signals, but signals are not received humans yet. So, this level deal before a human being can perceive the sound. At second, we can find physiological level. This level perceives sound based on the results of the extraction of various features. The sounds are implemented in the physiological response feature by the attribution of auditory perception and the nerve pathways, through the human senses. Next, we find psychological level. This level expresses or interprets the features of sound or music by classifying the weight of the features. Features are obtained at the physiological level and by its subsequent grouping. At last, cognitive level that interprets music by adopting a generic word (image word) for each of the groups classified at a psychological level.

When conducting experiments on music, it is very important to consider the source of sound to be investigated. However, we receive a different impression of musical performances given by world-class musicians from that of the automatic playing system. This difference can be explained by tempo fluctuations in musical performance (6).

3 Auditory Feature Values (7)

It can be considered that human beings take notice of some features of sound at a physiological level and can evaluate those features. We take notice of the following three points as the above features and seek the amount of features for the purpose of the human classification of the sound.

Pitch of sound is not measured directly and it is given by the attributes of auditory perception, which can perceive relatively the change of sound. Therefore, we consider it important to use the pitch of sound as feature values at a physiological level. And, Intensity of sound is difficult to measure directly the intensity of sound. Therefore, such a method is employed that measures relatively the intensity of sound by logarithmic, indicating a comparison of the intensity of two sounds. As the intensity of sound is one of the attributions of auditory perception, we consider it important to use it as feature values for the classification of the intensity of sound. Moreover, we take notice of the importance of time variation of sound. Most sounds have a time variation and it is important to perceive this time variation for the study of the auditory perception. As music is composed of the time variation of sound, we consider it important to use the time variation of sound as a feature value.

In this section, we aim to mention the feature value at a physiological, and to analyze the data with short-lasting Fourier transform. In section 4, we aim to model on a physiological level by conducting the retrieval experiment of similar music through the evaluation experiment of feature values.

3.1 Short-Lasting Fourier Transform of Music Data

A power spectrum can be obtained by the use of the short-lasting Fourier transform. In this study, we divided the result of the short-lasting Fourier transform into plural ranges by the following process in order to study the relationship between the time variation of sound and the attributes of auditory perception. Fig. 2 shows the result of the division. In Fig. 2, vertical axis represents the frequency period. Frequency period is divided by six ranges (Table.1). Horizontal axis represents the time period, is divided by “S” seconds. The value of “S” is established by the formula (1) with the following conditions.

A) Sampling frequency of the data used is 22050Hz.
B) Sampling frequency is halved in the process of the short-lasting Fourier transform.
C) In order to implement the short-lasting Fourier transform, the number of data should be 2 multiplied by n.

Table. 1 Range Division per frequency

	Range (Hz)
Treble	5000～10000
Middle Treble	2600～5000
Mediant	320～2600
Middle Bass	160～320
Bass	40～160
Bassy	20～40

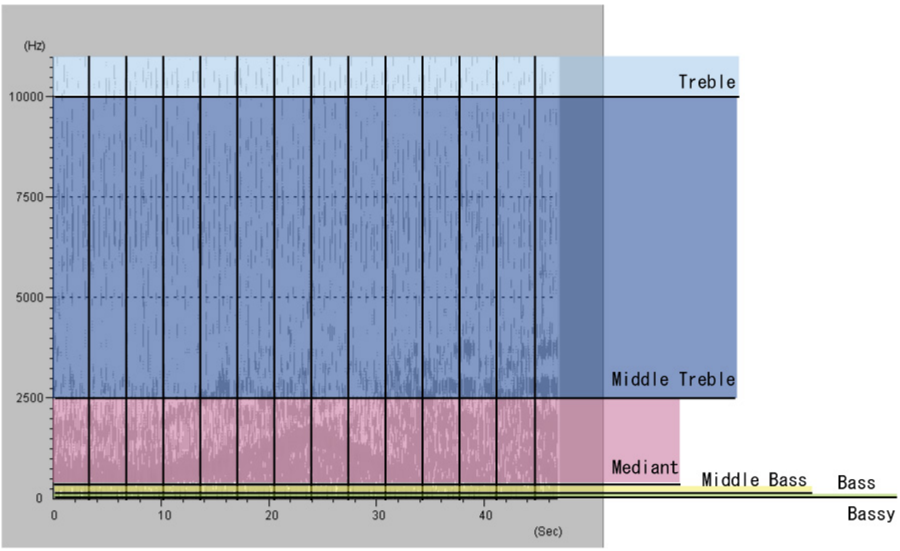


Fig. 2. Division of spectrum by range and time

$$s = \frac{2^{n}}{22050 \ / 2} = \frac{2^{12}}{11025} \approx 0.37 \ (n = 12) \ . \tag{1}$$

As it is set as "n=12", the time period per divided range is equal to the time per note of a tune with tempo of BPM 160. It is considered that the value of "n= 12" is the most appropriate for the time period per divided range. By performing the following processing of the divided ranges, the amount statistics is sought from feature values. For the next step, evaluation is made of this statistics quantity.

a) According to the Weber-Fechner Law, "The amount of perception is proportional to the logarithmic in the amount of stimulus," a logarithmic value (log PS) is sought of the power spectrum
b) Log PS is sought of each of the six (6) ranges divided.
c) The contrast of log PS is sought between the different ranges at the same time, "t."
d) The contrast of log PS is sought between the different ranges at different times, "t, and t+1."

In order to evaluate the significance of feature values and its statistics quantity, a retrieval experiment of similar music was performed by the use of the statistics quantity that was sought, based on the proposed feature values.

4 Experiment of the Similarity

The music data were collected from the 200 tunes of the RWC Music Database Collection (8). 10 tunes were selected at random among them, and the retrieval experiment of similar music was performed using these tunes as the key tunes. An evaluation was made of the proposed feature values by seeking the average relevance ratio of the top 5, 10, 15 and 20 similar music of the examination result. Further, in order to compare with the proposed feature values, the same experiment was performed using a power spectrum obtained from the fast Fourier transform as feature values.

The relevance ratios are shown in Table. 2 of the result of the retrieval experiment of similar music, by the fast Fourier transform and by the proposed feature values.

Table. 2 Relevance ratio of the top 20 similar music of the examination result

	~ 5th	～10th	~ 15th	~ 20th
Proposed feature values	72.8%	64.2%	56.4%	48.7%
Fast Fourier transform	59.1%	53.2%	47.5%	42.3%

The Wilcoxon rank-sum test was made using the similarity results and also an examination were performed of the significance of the difference of relevance ratios by the proposed feature values and by the use of the fast Fourier transform. As a result, at a significance level of 1%, the difference was recognized as significant between the two relevance ratios of the feature values.

5 Challenge and Perspective

In this study, we proposed the feature values of sound from the viewpoint of human attributes of auditory perception at a physiological level. With the results of the evaluation experiment of statistics quantity which was obtained from the proposed feature values, we have obtained the possibility of classification not depending on the name of tunes or composers, but by the use of the feature values which we proposed for this study, As a future policy, we will seek the improvement of statistics quantity, the most appropriate division of sound range, the increase in experiment data and others. Furthermore, by the use of the established auditory perception model, we are attempting the quantification / classification, by audio comparison experiment, of musical performances given by master musicians and by machines.

References

1. Kato, T., Sakai, K.: System/Information/Control. In: Research and Development of Sensitivity Agent and Human Media Data Base—Sensitivity Work Shop, vol. 42(5), pp. 253–259 (1998)
2. Kato, T.: Trans-category Retrieval Based on Subjective Perception Process Models. In: Proc. IEEE Multimedia and Expo ICME 2004, vol. TP9-5 (2004)
3. Ogino, A., Kato, T.: Design Approach of the Retrieval System of Sensitivity; Sensitivity System Modeling. Journal of the Information Processing Society of Japan; Data Base 46(SIG19(TOD 29)) (2005)
4. Tada, M., Kato, T.: Modeling of the Sensitivity of Auditory Perception by the Use of Hierarchical Classification and Application for the Retrieval of Similar Image. Journal of the Information Processing Society of Japan Data 44(SIG8(TOD 18)), 37–45 (2003)
5. Tada, M., Kato, T.: Feature Analysis of the Range of Similar Images and the Modeling of the Sensitivity of Auditory Perception. Journal of the Institute of Electronics, Information and Communication Engineers; D-II, J87-D-II(10), 1983–1995 (2004)
6. Ohmura, T.: Base of the Style of Musical Performance, pp. 158–165. Shunjusha Co., Ltd. (1998)
7. Moor, B.C.J., Ohkushi, K. (eds.) (Translation supervisor): Survey of Audio logical Psychology. Seisinshobo Co., Ltd. (1994)

Construction of a Model for Discriminating between Electroencephalographic Patterns at the Time of Incorrect Inputs Based on Sensitivity Spectrum Analysis

Raita Ohori, Daiki Shinkai, Yoshimitsu Nagai, and Syohei Ishizu

Department of Science and Engineering
Aoyama Gakuin University, Japan
C5610167@aoyama.jp

Abstract. We consider the possibility of prior discovery of slip-type human error by quantifying the "distraction" state that precedes it. We investigated brain activity during the task of keyboard-based text input taking incorrect input as a slip-type error. We used sensitivity spectrum analysis to quantify four output states. We attempted to detect in correct input by discriminating between the output of the distraction state before and after incorrect input and that of the state before correct input.

Keywords: Sensitivity spectrum analysis, electroencephalogram, event-related potential, distraction, error recognition, human error.

1 Introduction

Recently, when buying and selling stock, a certain company accidentally ordered "one share at 610,000 yen" instead of the intended "610,000 shares at 1 yen," and thereby incurred a substantial financial loss. Such failures are referred to as human error. Early detection of human error is important in preventing such losses. It is difficult to prevent all human errors, but measures should be taken to limit them. Examples of measures intended to limit human error are records management and fool proofing. Records management allows human error to be tracked by recording and accumulating data, and can prevent the re-occurrence of an error by information sharing. Fool proofing entails assuming that any possible error will happen, and taking measures to avoid problems after an error occur. In either case, however, human error is not prevented from happening beforehand.

According to British cognitive psychologist J. Reason, there are two types of human error: planning failures and practice failures. The former is called a slip or a lapse, and the latter is called a mistake. In the case of a slip, the plan itself is correct, but there is a failure in its implementation. A lapse is an error in which the plan is forgotten during the course of implementation. In a mistake, implementation of the intended plan is correct, but the plan itself is flawed. Slips in particular are related to failures of attention, and these are the only errors in which action is routine and there is no consciousness of the correctness of the plan. In other words, there is a failure to pay sufficient attention, for instance, due to fatigue or worry, and the actor enters a

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 618–626, 2011.

"distraction" state that causes the error. We therefore wondered whether it might be possible to detect slips beforehand by quantifying the distraction state.

It is known that monitoring the activity of the human brain can reveal changes in mental state, such as entering a distraction state. Musha [1] proposed using electroencephalography to quantify the sensitivity of mental states such as "anger", "satisfaction", "sorrow", and "pleasure". This technique is called the sensitivity spectrum analysis method. Because mental states are complex, it is difficult to directly specify the relationship between active parts of the brain, even using information on the occurrence or non-occurrence of brain waves, such as alpha and beta waves. Sensitivity spectrum analysis uses three elements—time, frequency, and space—to quantify a sensitivity state by assembling a matrix that assumes feelings as an output vector. By replacing the four states of "anger", "satisfaction", "sorrow", and "pleasure" with states pertaining to the occurrence of error, we hoped that prior discovery of slip-type human error would be possible if we could assume a distraction state in the output.

In this study, we used keyboard-based text input into a PC as a task for investigating the distraction state. It is common for slips to occur when performing such a task. As performed by an expert, text input is largely automated: in cases of incorrect input, the plan for inputting is itself correct, but at some stage, a wrong key is typed and thus failure of implementation occurs. In other words, the incorrect input is slip-type human error. We consider a keyboard input task and use electroencephalography to quantify the distraction state before the incorrect input, which is a slip-type error. By quantifying the distraction state, we can expect prior discovery of slip-type human errors

The purpose of this study is quantification of the distraction state that precedes slip-type human errors. We used a PC-based text input task and considered incorrect input as a slip. We used sensitivity spectrum analysis to quantify the mental states. We converted electroencephalographic data into four output states of inputs and attempted to detect incorrect inputs by discriminating between the output of the distraction state before and after incorrect input and that of the state before and after correct input.

2 Electroencephalography and Sensitivity Spectrum Analysis

2.1 Electroencephalography

Electrical activity in the brain is measured as brain waves. Here, we describe frequencies and event-related potentials that are characteristic reactions, and their measurement values.

An electroencephalogram contains delta waves, theta waves, alpha waves, and beta waves, which are classified according to their frequency:

Delta waves (0.5–4 Hz)	Sound sleep
Theta waves (4–8 Hz)	Drowsiness
Alpha waves (8–13 Hz)	Mental rest
Beta waves (13–40 Hz)	Mental activity

2.2 Event-related Potential (P300)

A limitation of electroencephalographic analysis is that signals resulting from a given stimulation of a specific sense are weak. The addition average law is used to overcome for this. The addition average law is applied to electroencephalography by performing the same stimulation tens of times, recording the electroencephalogram each time, and matching recorded data from the times of stimulations to add and average the signals. Waves irrelevant to the target stimulation will have different measurements each time, and so eventually will be averaged out and become approximately flat. Confirmed changes in the electroencephalogram are called event-related potentials. Regardless of the type of stimulation, significant changes occurring 0.3 s before or after are called P300. It is said that the P3, Pz, and P4 spots of the international 10-20 system can confirm many P300 waveforms.

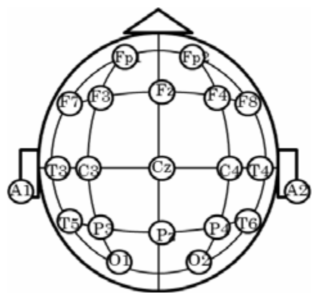


Fig. 1. Electroencephalogram measurement positions of the international 10-20 system

2.3 Sensitivity Spectrum Analysis

We shall next discuss the principles of sensitivity spectrum analysis. Ten electrodes are arranged on the scalp, and the electroencephalogram is recorded. The equation in Formula (1) is then used to find the cross-correlation coefficient between the electrodes. Cross-correlation coefficients have the following advantages:

1. They are not influenced by individual differences in magnitude between electroencephalogram amplitudes.
2. They can express phases between the parts of the electroencephalogram, such as polarity relations.
3. They can express both wave pattern information and information between parts at the same time.

$$X_{i,j,k} = \frac{\sum (X_{i,k}(t) X_{j,k}(t))}{\sqrt{\sum (X^2_{i,k}(t)) \sum (X^2_{j,k}(t))}} \quad (1)$$

In Formula (1), $X_{i,j,k}$ are the cross-correlation coefficients, $i, j = 1 \sim 10$ are the measurement points, $k = \theta, \alpha, \beta$ are the three types of brain waves, and t is time.

For 10 chronologically ordered measurements, 45 cross-correlation coefficients are obtained and 135 variables are obtained when every frequency is needed. The 135 variables, arranged in order, are called a state vector. Four states of feelings are assigned values of 0 or 1, and four variables to express those levels are called a sensitivity vector. Linear conversion is used to derive a sensitivity vector from the state vector and a 135×4 matrix is used for the conversion. This is called a sensitivity matrix (Fig. 2).

$$\underbrace{\begin{pmatrix} a_{1,1}, a_{1,2} \ldots a_{1,135} \\ a_{2,1}, a_{2,2} \ldots a_{2,135} \\ a_{3,1}, a_{3,2} \ldots a_{3,135} \\ a_{4,1}, a_{4,2} \ldots a_{4,135} \end{pmatrix}}_{\text{Sensitivity matrix}} \times \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{135} \end{pmatrix}}_{\text{State vector}} + \underbrace{\begin{pmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \end{pmatrix}}_{\text{Constant term}} = \underbrace{\begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{pmatrix}}_{\text{Sensitivity vector}}$$

Fig. 2. The sensitivity spectrum analysis model

In this study, we use sensitivity spectrum analysis to quantify the output of the distraction state before incorrect input and to quantify the output of other states before correct input in a keyboard-based text input task.

3 Creating the Incorrect Input Sensitivity Matrix

3.1 Electroencephalogram Measurement Experiment

Experiment Preparations. We used a non-adhesive electrode helmet (BFL-EH15) to record the electroencephalogram. Nine electrode positions were used: Fp1, Fp2, F3, F4, P3, P4, O1, O2, and Pz, as per the locations shown in Fig. 1. We used the right earlobe as a standard electrode, and conducted the measurement using monopole induction.

Signals sent from the electrodes underwent A/D conversion, and a processor box recorded, digitized, and filtered the data. We transferred output data for storage and analysis on a computer through a SCSI bus interface. During the experiment, we displayed words for keyboard-based text input on a computer in a soundproofed room, and electroencephalograms were measured at the time of incorrect inputs. The screen on the computer used for text input was partitioned into two areas. The lower partition displayed the sentence to be typed, and the upper partition was used for text input.

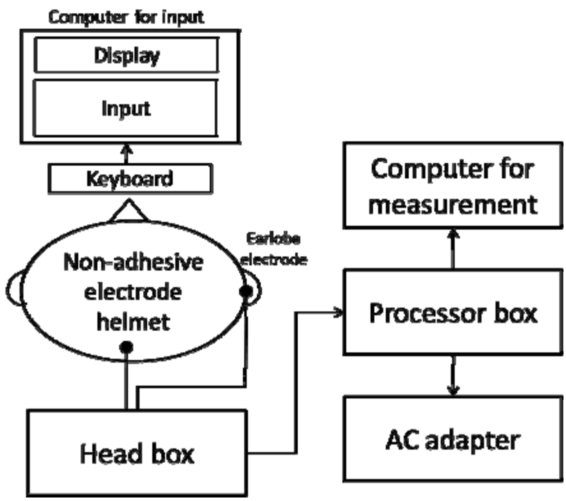


Fig. 3. Experiment overview

Experimental Procedure. The participants were nine healthy men aged 21–24 years old. We explained the text input measurement to them without mention of the incorrect input measurement, and after obtaining their consent, conducted the test according to the procedure given by steps 1–4 below.

1. Participants sat in front of the text input computer and an electroencephalogram was recorded.
2. Participants closed their eyes and rested for 1 min.
3. Each sentence consisting of around 100 kanji and hiragana characters was shown in the display partition of the computer screen.
4. Participants used the keyboard to input into the input partition the sentence displayed in the display partition.

We measured electroencephalograms for 5 types of sentences for each participant, with breaks of 1 min between inputting sentences.

Results of the Experiment. Table 1 shows the number of incorrect inputs from the nine participants (A–I).

Table 1. Number of incorrect inputs for each participant

Participant	A	B	C	D	E	F	G	H	I	Total
Incorrect inputs	15	14	9	24	21	18	16	14	21	152

Separately from data related to incorrect inputs, we also recorded data related to correct inputs. Funada has confirmed that P300 waveforms appear as a result of keyboard input processes [2]. In this knowledge of P300, we consider that a change can be seen in electroencephalograms approximately 0.3 s after an incorrect input. The time duration before occurrence of the waveform was taken to be 0.2 s, because the average time before hitting the backspace after an error was 0.4 s. We classified electroencephalographic recording before incorrect input as that related to the distraction state. We took the time duration of distraction to be 1.2 s (from 1.0 s before the incorrect input to 0.2 s after the incorrect input).We used the same time duration for the correct input state. We found cross-correlation coefficients between the electrodes to use in determination of the time duration. We found cross-correlation coefficients at the times of incorrect inputs and correct inputs in the distraction state and in the error recognition state during the time duration, and developed the sensitivity matrix under the assumption that we could detect in correct input from the relationship between the distraction state and the matrix value. In other words, if the discrimination precision of the matrix is high, the matrix value of the distraction state will rise before the incorrect input, and the matrix value of the error recognition state should rise after incorrect input. We show the time duration of the incorrect input in Fig. 4.

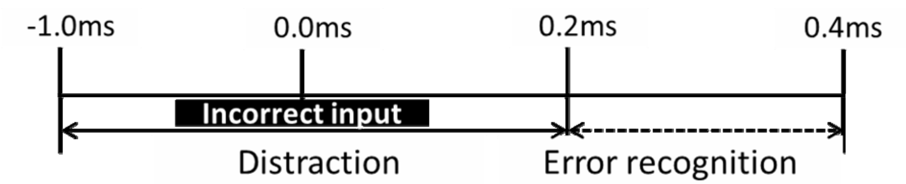


Fig. 4. The distraction state and the error recognition state in the time duration

3.2 Creating the Incorrect Input Sensitivity Matrix

Cross-correlation Coefficient. We recorded electroencephalograms showing theta waves, alpha waves, and beta waves from nine measurement points during each time duration. We found 108 cross-correlation coefficients to interlace with provided data using the equation shown in Formula(1).

Next, in the cross-correlation coefficient of the electroencephalographic data for the error recognition state and the normal input state, we show pairs of significant differences between averages on the horizontal axis (Fig. 5). In Fig. 5, we found that theta waves had many pairs of significant differences in the cross-correlation coefficient. In particular, the F3 measurement point showed the largest number of significant differences. Significant differences were also seen in Pz. We therefore had found a difference in the cross-correlation coefficient, and so considered the cross-correlation coefficient obtained when creating the sensitivity matrix to be useful.

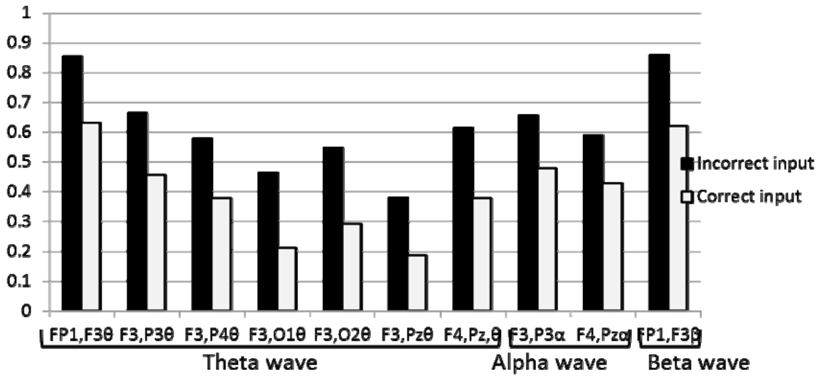


Fig. 5. The cross-correlation coefficient of electroencephalogram data of the incorrect input state and the correct input state

Creating the Incorrect Input Sensitivity Matrix. For the sensitivity matrix, we derived a matrix to output the state of error recognition and the state after normal input, and a matrix to output the state of distraction and the state before the normal input. We used the sensitivity spectrum analysis model of Fig. 2, creating a sensitivity spectrum analysis model that output data for times of incorrect input and correct input, since in this study we used the data from times of correct and incorrect input as substitutions for the four states of feelings (Fig. 6). The sensitivity vector before the time of the incorrect input outputs Z_1, Z_2. And, the sensitivity vector after the time of the incorrect input outputs Z_1', Z_2'.

$$\underbrace{\begin{pmatrix} a_{1,1}, a_{1,2} \ldots a_{1,108} \\ a_{2,1}, a_{2,2} \ldots a_{2,108} \end{pmatrix}}_{\text{Sensitivity matrix of Incorrect input}} \times \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{108} \end{pmatrix}}_{\text{State vector}} + \underbrace{\begin{pmatrix} C_1 \\ C_2 \end{pmatrix}}_{\text{Constant term}} = \underbrace{\begin{pmatrix} z_1 \\ z_2 \end{pmatrix}}_{\text{Sensitivity vector}}$$

Fig. 6. Sensitivity spectrum analysis model at the time of the incorrect input

We assumed the cross-correlation coefficient that we found with the state vector. We used 1 as the sensitivity vector in each expression, and 0 for other values. We found the 108 coefficients by using multiple regression analysis in order to find the sensitivity matrix. We assumed the intercept found from multiple regression analysis as the constant term. In Table 2, a_1 is the matrix of fixed number values before incorrect input (the distraction state), a_2 is the values before correct input, a_1' is the values after incorrect input (the error recognition state), and a_2' is the values after the correct input.

Table 2. Constant values of the incorrect input sensitivity matrix

	1	2	…	107	108
a_1	-1.9595	4.9017	…	-7.5581	1.8951
a_2	7.4453	-76.8792	…	12.9451	11.0375
a_1’	2.2452	14.4860	…	-2.9389	-2.6061
a_2’	1.5692	-2.3094	…	1.4934	1.3092

Verification of Discrimination Precision. We verified the discrimination precision of the constructed matrix. We set Z_1 as the sensitivity vector value before an incorrect input (the distraction state), and Z_2 is the value before a correct input. Where $Z_1 - Z_2 > 0$, the state is before an incorrect input (the distraction state). Where $Z_1 - Z_2 < 0$, the state is before a correct input. Likewise, we set Z_1' as the sensitivity vector value after the incorrect input (the error recognition state), and Z_2' as the value after correct input. We had a participant input sentences and acquired 10,000 data points (we measured brain waves every 1/100 s) in 100 s. We detected incorrect input with these data during the time duration (0–120, 1–121…9880–10000) by a comparison with the matrix before the input. Likewise, we detected the incorrect input with these data during the time duration (0–20, 1–21…9980–10000) by a comparison with the matrix after the input. We found the chronological order change of $Z_1 - Z_2$, $Z_1' - Z_2'$ before and after the incorrect input, and we set the vertical axis as vector values, and the horizontal axis as time (Fig.7). Table 3 shows the discrimination precision results for all data.

Table 3. The discrimination precision of the sensitivity matrix for incorrect input

	Before incorrect input	Before correct input	After incorrect input	After correct input
Total data	534	4890	534	4890
Discriminated data	434	2542	427	3032
Discrimination rate	88%	52%	83%	62%

As shown in Fig. 7, the incorrect input is performed at 21 s (indicated by the double circle).The vector value of $Z_1 - Z_2$ is large before the incorrect input (at 0.31 s), and the vector value of $Z_1' - Z_2'$ becomes large after the incorrect input; thus, we could discriminate between states before and after incorrect input. Table 3 shows that, for all data, 88% and 83% of states respectively before and after incorrect input could be discriminated. However, the discrimination of states before and after correct input was 52% and 62%, respectively. We attribute this discrepancy to the difficulty of treating data as equivalent despite influences such as differences between input words and differences in what is being thought at the time of the typing input. We were, however, able to quantify the state at the time of incorrect input using the sensitivity matrix.

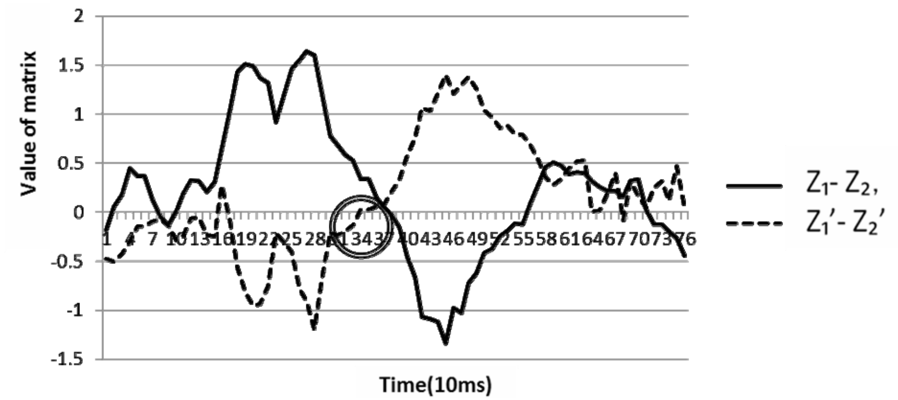


Fig. 7. The chronological order change of the sensitivity vector at the time of an incorrect input

4 Conclusion

In this study, we quantified the "distraction" state that precedes slip-type human error. We used a keyboard-based text input task and assumed incorrect input to be a slip. We examined four output states of inputs, using electroencephalography as a technique for quantifying the state, and using sensitivity spectrum analysis. We found that, with states of inputs output, it was possible to detect incorrect inputs by discriminating between the distraction state that precedes incorrect input and the state before correct input.

References

1. Araga, Y., Musha, T.: Numeric quantification of human emotions with EEGs. Bioscience &Industry 59(10), 37–38 (2001)
2. Funada, M., Ninomiya, S.: A Discrimination Algorithm of Event Related Potentials Evoked by Characters. BioMed. (2004)
3. shimizu, T., Nakagawa, M.: Measurement of visual and sensational information by Images with Fractal Analyses of EEG. The Institute of Electronics, Information and Communication Engineers 108(219), 97–102 (2008)

Basic Study of Analysis of Human Brain Activities during Car Driving

Noboru Takahashi[1], Shunji Shimizu[2], Yukihiro Hirata[2], Hiroyuki Nara[3], Hiroaki Inoue[1], Nobuhide Hirai[4], Senichiro Kikuchi[4], Eiju Watanabe[4], and Satoshi Kato[4]

[1] Graduate School, Tokyo University of Science, Suwa, Japan
[2] Tokyo University of Science, Suwa, Japan
[3] Hokkaidou University, Japan
[4] Jichi Medical University, Japan
jgh10614@ed.suwa.tus.ac.jp

Abstract. Recently, as the worldwide population grows older, it is thought that various assistive systems have soared. It is necessary to discuss functions of humans, spatial perception, decision-making, and determining direction, for developing assistive systems. The final goal of our research is to contribute to developing of welfare robots with functions that are responsive like human. We measured brain activities during virtual car driving using NIRS. As a result of these experiments, there were significant differences at outside frontal cortex in left brain. This go-round, we measured brain activity during actual car driving. In general roads, experiments were performed by taking f-NIRS in the car, and measuring the brain activity when car driven by subjects was went through a number of intersections and road signs. In addition, there was significant difference in common regions. We report a significant association between car driving and brain activity.

1 Introduction

Human movements change relative to his environment. Nevertheless, he recognizes a new location and decides what behavior to take. It is important to analyze the human spatial perception for developing autonomous robots or automatic driving.

The relation of the theta brain waves to the human spatial perception was discussed in [1] [2]. When humans perceive space, for example, try to decide the next action in a maze, the theta brain waves saliently appear. This means we have a searching behavior to find a goal at an unknown maze. From the side of human navigation E.A.Maguire et.at.al measures the brain activations using complex virtual reality town [3]. But each and every task is notional and the particulars about the mechanism that enables humans to perceive space and direction is yet unknown.

Recently, functional localization of the brain has been gradually clarified through many researches on brain functions. It is well known that in the frontal lobe higher order processing is done such as of memory, judgment, reasoning, etc. However, there is little information about what happens in the frontal lobe when every act of driving taken. We grasp mechanism of information processing of the brain by analyzing data

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 627–635, 2011.

on human brain activity during car driving[4][5]. The goal of this study is to find a way to apply this result to new assist system with human motions. To achieve the goal, the brain activity of frontal lobe, which related to behavioral decision-making is discussed from the viewpoint of human spatial perception. In particular, a driving movie is shown to the subjects as sensory information. We measured the brain activity of frontal lobe how concern to the handling a wheel motion of drivers connects in action of car driving. Furthermore, the brain activity is to be measured and discuss the mechanism of information processing of the brain by analyzing experimental data concerning human brain activity during car driving using NIRS.

2 Experiment Method

2.1 Brain Activity on Virtual Driving

At the first experiment, we measured the density of oxygenated hemoglobin (oxy hemoglobin) and deoxygenated hemoglobin (de oxy hemoglobin) in the frontal cortex area while subjects were shown the driving movie. An NIRS (Hitachi Medical Corp ETG-100) with 24 channels (sampling frequency 10 Hz) was used to record it.

The subjects for this experiment were eight males aged 22 to 24. The average age was 22.7 and the age of standard deviation was 0.74. All of the subjects were right handed. They were asked to read and sign an informed consent regarding the experiment. Subjects were seated in car seat. Then they were fitted with the NIRS probe and the HMD. They were covered with black cloth to shut out the light from outside.

Driving movie for the experiment was recorded from a car with a video camera aimed toward the direction of movement. The movie is included two scenes at a T-junction in which it must be decided either to turn to the right or left(fig1). In the second scene, there is a road sign with directions. We used nine kinds of movies in about one minutes.

Before showing the movie, subjects were given directions to turn to the right or left at the first T-junction. They were also taught the place which was on the road sign at the second T-junction. They had to decide the direction when they looked at the road sign. They were asked to push a button when they realized the direction in which they were to turn.

Subjects take during 10 seconds with their eye close before movies were shown and they view the image after that. Then the brain activity was recorded from the first eyes-closed rest to the last eyes- close rest.

Here we will define Tasks A, B, and C; Tasks A and C were proposed as the same experiment tasks. After the subjects had pushed the button, other operation was added in Task B. It is the operation that the steering wheel was turned in the direction of destination when the subjects were able to judge the task proposed.

At second experiment, measuring was peformed by f-NIRS(Functional Near Infrared Spectroscopy) made by SHIMADZE Co. Ltd with 44ch Five subjects were a healthy male in their 20s, right handed with a good driving history.

The subject was asked to perform simulated car driving, moving their hand in circles as if using a steering wheel. A PC mouse on the table was used to simulate handling a wheel, and NIRS(near-infrared spectroscopy) to monitor oxygen content change in the subjects' brain. NIRS irradiation was performed to measure brain activities when the subject sitting on a chair make a drawing circle line of the right/left hand 1) clockwise, and 2) counterclockwise.The part of measurement was the frontal lobe. The subject was asked to draw on the table a circle 30 cm in diameter five times consecutively, spending four seconds per a circle. The time design was rest(10 seconds at least) -task(20 seconds) - rest(10 seconds).

Fig. 1. Two T-junction included in driving movie

2.2 Brain Activity During Actual Car Driving

In general roads, experiments were performed by taking f-NIRS in the car, and measuring the brain activity when car driven by subjects was went through two different intersections. Six subjects were a healthy male in their 20s, right handed with a good driving history.

Subjects close their eyes for 10 seconds at least, and drove the car for 600 seconds. Subjects were given directions to turn to the right or left at the first T-junction during 600 seconds. They were also taught the place which was on the road sign at the second T-junction. And, they were given the place where they have to go to. So, they had to decide the direction when they looked at the road sign during 600 seconds. Finally, subject close their eyes for 10 seconds again and become the end. 3 pattern were prepared for the task pattern.

2.3 Verification Experiment

To conduct verification for experimental results in actual driving, we performed additional experiment which was achieved in a similar way. In this experiment, experimental course was different from previous experiment. While previous one was included two T-junctions in which there was road sign at second one and not at first one per a measurement, there were multiple T-junctions. Three were 5 T-junction without road sign and 4 T-junctions with road sing. Subjects were twelve males who were all right-handed.

3 Result for the Experiment

3.1 Brain Activity on Virtual Driving

For task A and B in the first experiment, the suggested movies the subjects were headed to informed direction, and they let decided which way to turn under the road sign. At the T-junction, they were to push the button when they realized the direction. In task A, the task was added after the button indicating the direction to turn was pushed by the subjects. The hemoglobin variation was compared in the results of Tasks A and B, A and C to see the brain activity pertaining to special perception during the same movie.

Equation (1) was used to compare the data. τ_1 was set the time as its length was 1 second before being pushed the button. Similarly, τ_2 was set in a way similar to τ_1. And x_i (t) indicates variation of i channel oxy hemoglobin or deoxy hemoglobin. We then took a average of $x_i(t)$ throughτ_1 andτ_2. In this situation, i of the defined c (i) was the channel for the brain activity. Because of the sampling frequency was set on 10 Hz, we calculated 10 times per sec. Fig 2 shows the sample result for this calculation c(i) in task A with oxy-hemoglobin(ch13 was zero).

$$c(i) = \overline{\sum_{\tau_2} x_i(\tau_2)} - \overline{\sum_{\tau_1} x_i(\tau_1)} \qquad (1)$$

The next step was to calculate the average of all subjects to compare the situations in which the steering wheel was turned and when it was not.. Fig 3 and 4 show the results. Many upward tendencies could be found in fig 3. This might have occurred

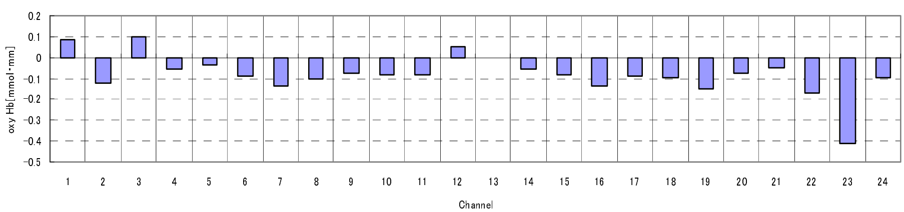


Fig. 2. Result from subject A of task A

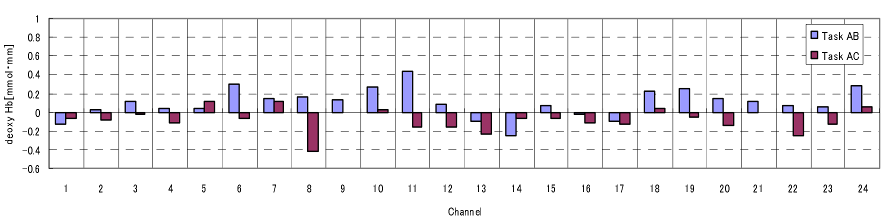


Fig. 3. Compared between turning the steering wheel and not (oxy Hb of average)

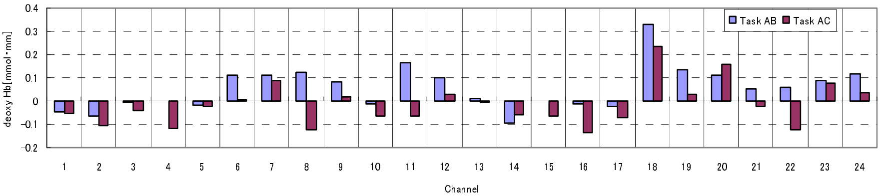


Fig. 4. Compared between turning the steering wheel and not (deoxy Hb of average)

when they were finding the way from a road sign. In addition, the results indicate a greater increase when the subjects turned the steering wheel. That means observation of brain activity has been made during movement based on spatial perceptions.

On the whole, the variation in de-oxy hemoglobin (fig 4) was smaller than in the oxy hemoglobin. However, there was a great increase in Channel 18. This might be the variation based on the spatial perceptions.

Calculation result of him with the tasks A and C. Next, differences were investigated concerning the subject's brain activity. As the First case, it was when the vision was directed after having been told the direction. As the Second, it was when the vision was directed after having been decided the direction under the road sign.

d_1 and d_2 shown in fig 5 are defined as below. d_1 is the variation ofhemoglobin turning at the first T-junction. And d_2 is variation of hemoglobin at the second one. From the measurement result, d_1 and d_2, all of the 269 times of each subject, there were significant differences in oxy hemoglobin 3ch. ($p<0.02$: paired t test) and 20ch. ($p<0.03$) in fig 6 using NIRS.

Subjects pushed a button before turning at the second T-junction, so it influenced brain activities. The possibility of a correlation between d_2 and the time until the movie was turned at the second T-junction after each subject pushed a button was investigated. Each correlation coefficient of hemoglobin channel was calculated.

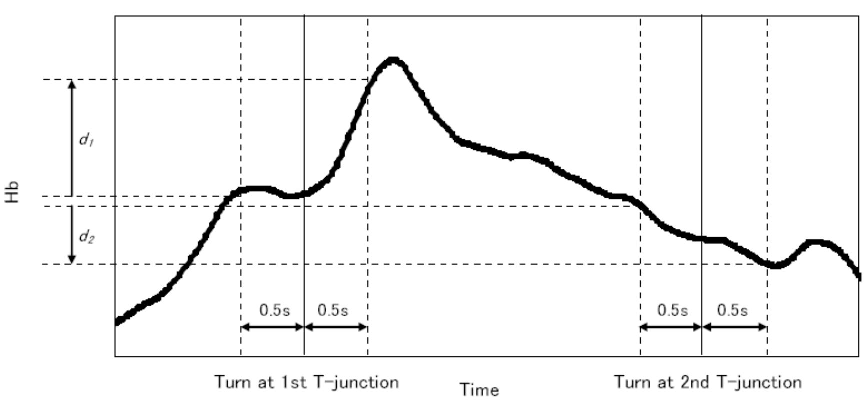


Fig. 5. Define variation of hemoglobin d1 and d2

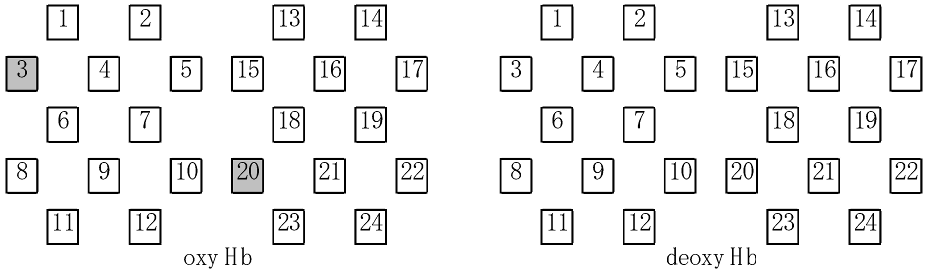


Fig. 6. Significant differences at NIRS's oxy-hemoglobin are gray

There was significant difference at only de-oxy hemoglobin 10ch. ($p<0.07$) using paired t test. In only this result, the relationship between pushing a button and d_2 cannot be judged.

During the motion in the second experiment, the increase of oxy hemoglobin density of the brain was found in all subjects(fig7). The different regions of the brain were observed to be active, depending on the individual.

The subjects were to be observed 1) on starting, and 2) 3-5 seconds after starting moving their 3) right hand 4) left hand 5)clockwise 6)counterclockwise. Although some individual variation existed, the result showed the significant differences and some characteristic patterns.

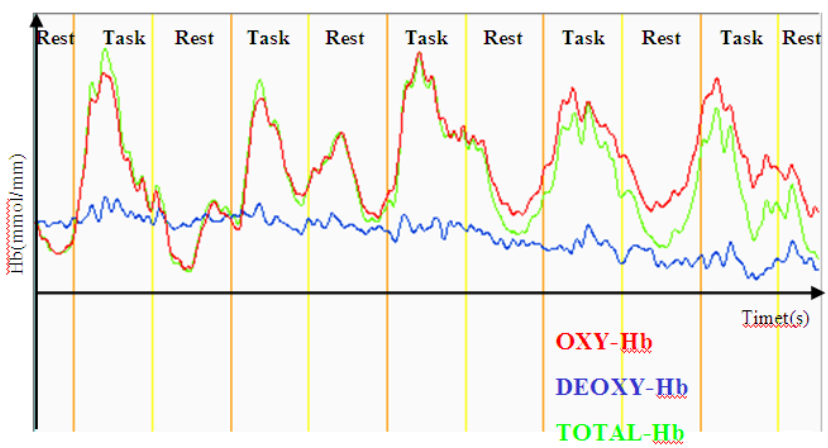


Fig. 7. The brain activity of some subject

The obtained patterns are shown as follows. Regardless of 1), 2), 3) and 4) above, the change in the oxy hemoglobin density of the brain was seen within the significant difference level 5% or less in the three individuals out of all five subjects. The part was the adjacent part both of left pre-motor area and of left prefrontal cortex. Especially, in the adjacent part of prefrontal cortex a number of significant differences were seen among in four out of five subjects.

Next more emphasis was put on the rotation direction: 5) clockwise or 6) counterclockwise. No large density change was found in the brain with all the subjects employing 6). But the significant difference was seen in four out of five subjects employing 5). (fig.8) It is well known that in the outside prefrontal cortex higher order processing is done such as of behavior control. It is inferred that the pre-motor area was activated when the subjects moved the hand in the way stated above because the pre-motor area is responsible for behavior control, for transforming visual information, and for generating neural impulses controlling.

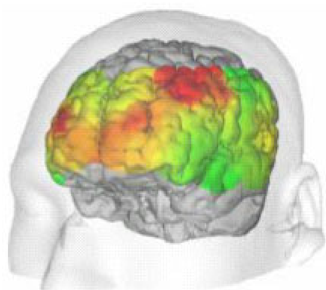


Fig. 8. Brain activity (clockwise)

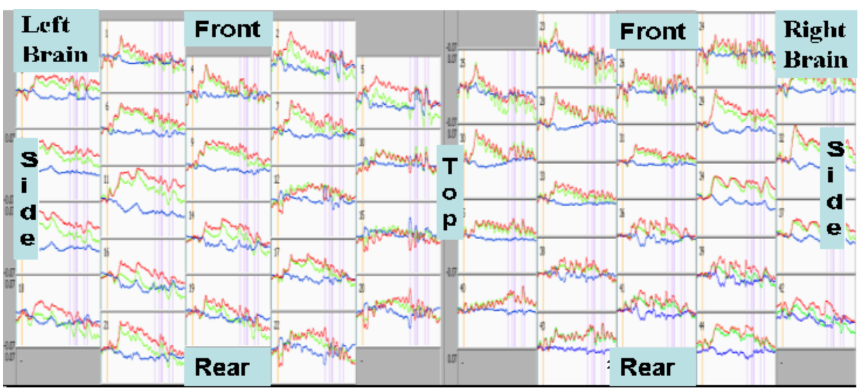


Fig. 9. The brain activity of some subject

3.2 Brain activity on Car Driving

At the first, Hb-oxy was increased in overall frontal lobe after start of operation. This tendency is common among subjects. After that, Hb-oxy was decreased as subjects adjusted to driving the car. This meant that the brain activity changed from collective to local activities. Fig 9 was shown one subject's brain activity. In this experiment, being considered time as zero when subjects turned a steering wheel. The analysis was performed one-sample t-test within the significant difference level 5% or less between zero and about four seconds after turning . As the results, there were significant differences around #46 area and #9 area of the dorsolateral prefrontal cortex and the premotor area of the left hemisphere brain at the turn left. Around #46 area is corresponded to working memory.

3.3 Verification Experiment

Analysis method was the same as previous one. Though Gaps were shown regions at which there were significant, there were significant differences in common region, too. It was thought that there were significant differences at difference and common regions because of different task.

4 Conclusion

The hemoglobin density change of the human subjects' frontal lobe is partly observed in the experiments we designed, where three kinds of tasks were performed to analyze human brain activity from the view point of spatial perception.

The NIRS measures of hemoglobin variation in the channels suggest that human behavioral decision-making of different types may cause different brain activities as we saw in the tasks: 1) take a given direction at the first T-junction, 2) take a self-chosen direction on a road sign at the second T-junction and 3) turn the wheel or not. Some significant differences (paired t test) on NIRS's oxy-hemoglobin and less interrelated results between "pushing a button" and brain activity at the second T-junction are obtained. Researches into other human brain activities than spatial perception are to be necessary with accumulated data from fMRI, EEG,etc.

Furthermore, experimental results indicate that with the subjects moving their hand in circle, regardless of right or left, 1) the same response was observed in the prefrontal cortex and premotor area, and 2) different patterns of brain activities generated by moving either hand clockwise or counterclockwise.

The regions observed were only those with the 5% and less significance level. Possible extensions could be applied to other regions with the 10% and less significance level for the future study. With a larger number of subjects, brain activity patterns need to be made clear.

In addition, it is thought to take particular note of participation concerning working memory when car is driven. From Results of these experiments, there was significant difference around working memory. So, experiments focusing on relationship turning wheel and working memory will be performed.

Acknowledgement. Everyone in HITACHI Co. Ltd, SHIMADZU Co. Ltd and Shinapex Co. Ltd are profoundly acknowledged. The study was supported by them.

References

1. Kahana, M.J., Sekuler, R., Caplan, J.B., Kirschen, M., Madsen, J.R.: Human theta oscillations exhibit task dependence during virtual maze navigation. Nature 399, 781–784 (1999)

2. Nishiyama, N., Yamaguchi, Y.: Human EEG theta in the spatial recognition task. In: Proceedings of 5th World Multiconf. On Systemics, Cybernetics and Informatics (SCI 2001), 7th Int. Conf. On Information Systems, Analysis and Synthesis (ISAS 2001), pp. 497–500 (2001)
3. Maguire, E.A., Burgess, N., Donnett, J.G., Frackowiak, R.S.J., Frith, C.D., O'Keefe, J.: Knowing Where and Getting There: A Human Navigation Network. Science 280 (May 8, 1998)
4. Shimizu, S., Hirai, N., Miwakeichi, F., et al.: Fundamental Study for Relationship between Cognitivetask and Brain Activity during Car Driving. In: The 13th International Conference on Human-Computer Interaction, San Diego, CA, USA, pp. 434–440. Springer, Heidelberg (2009)

Bereitschaftspotential Modeling by DBNM and Its Application to BCI

Shino Takata[1], Toshimasa Yamazaki[1], Maiko Sakamoto[1], Takahiro Yamanoi[2], and Kenichi Kamijo[3]

[1] Kyushu Institute of Technology, Japan
[2] Hokkai Gakuen University, Japan
[3] NEC Green Innovation Research Labs, Japan
kingsma51@yahoo.co.jp, t-ymzk@bio.kyutech.ac.jp

Abstract. In this study, movement-related potentials (MRPs) including Bereitschaftspotential (BP) is modeled by a dynamic Bayesian network model (DBNM). The containing the MRPs BP are divided into the early BP, NS' (negative slope) and MP (motor potential) intervals, each of which is represented by a Bayesian network model (BNM). Each BNM is constructed using the results by equivalent current dipole source localization (ECDL) after independent component analysis (ICA), for single-trial EEGs recorded during the hand movement. Nodes in the BNM correspond to the brain sites where dipoles are located. Connecting with the three BNMs yields a DBNM. This model is used to discriminate between the left- and right-hand movements in a framework of single-trial-electroencephalogram (EEG)-based BCI.

1 Introduction

Brain-Computer Interfaces (BCIs) refer to the technology for translating users' brain activities, especially scalp-recorded electroencephalogram (EEG), into their intention and for transporting to the external world [1].

In a framework of single-trial-EEG-based BCI, some of the authors had already constructed a Bayesian network model (BNM) to discriminate between left and right hands to be imagined grasping from single-trial EEGs recorded during the movement imagery tasks [2],[3]. Taking time course of the brain activities into account, however, we should generalize the BNM to a dynamic BNM (DBNM), and the latter could improve the performance of the former. In order to validate this improvement, in this study, movement-related potentials (MRPs) including Bereitschaftspotential (BP) are modeled by a DBNM. The MRPs containing the BP are divided into the early BP, NS' (negative slope) and MP (motor potential) intervals, each of which is represented by a BNM. Each BNM is constructed using the results by equivalent current dipole source localization (ECDL) after independent component analysis (ICA), for single-trial EEGs recorded during the hand movement. Nodes in the BNM correspond to the brain sites where dipoles are located. *Node activities* are assumed to be conditional probabilities in the nodes. Connecting with the three BNMs through each of the most superior nodes yields a DBNM. This model is used to discriminate between the left- and right-hand movements.

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 636–640, 2011.

2 Methods

Figure 1 shows an overview of the present experimental system. 32-channel single-trial EEGs are recorded during the following tasks. For any one trial, on the display ”choice”, “+” 3.5 s after “choice” and “report” 2.0 s after “+” in turn are presented, whose durations are 2.0 s, 0.2 s and 2.5 s, respectively. In case of “left” or “right” selection in “choice”, the subject is requested to grasp his hand of the same side as the selection as soon as possible when “+” is presented, and then to speak aloud the selection in “report”, otherwise no grasping and no speaking(Fig.2). During this task, 32-channel single-trial EEGs, electromyogram (EMGs) and electrooculogram (EOGs) are recorded, where the EOGs are used to monitoring eye movements [4].

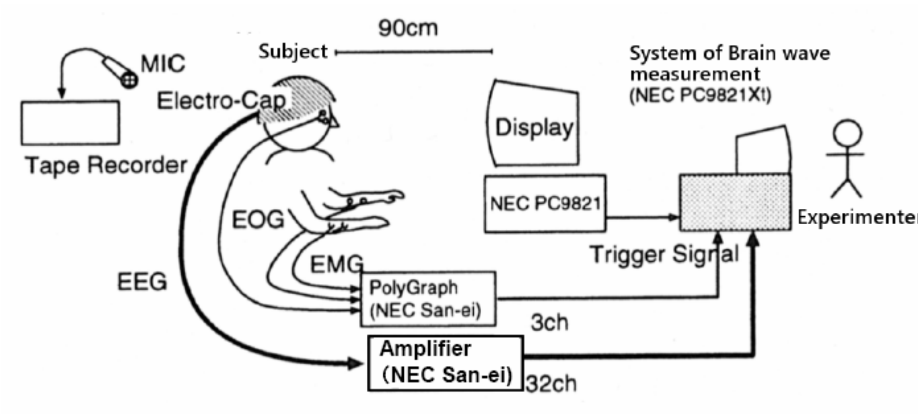


Fig. 1. Configuration of experimental system

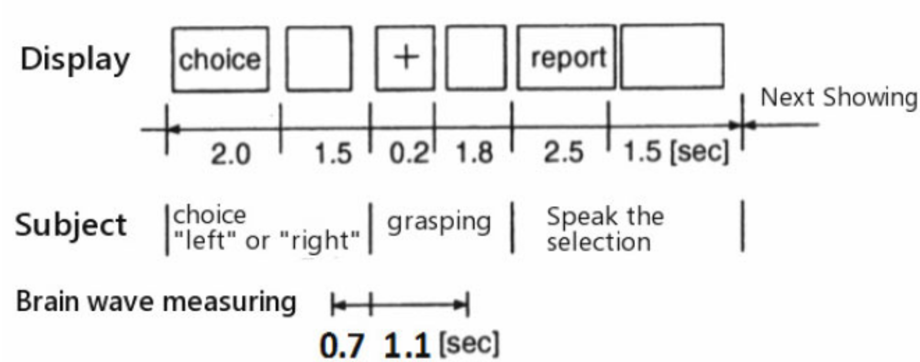


Fig. 2. Experimental procedure

Assuming that the onset of the EMG is set to be 0 ms, the early BP (E-BP) refers to the interval between -1700 and -400 ms, NS’ that between -400 and 0 ms, MP that between 0 and 200 ms and "other" after 200 ms, for each single-trial EEG. ICA [5] is applied to the 32-channel EEGs for each interval, and then the ECDL to the reconstructed EEGs after the deflation procedure, using "SynaCenterPro" (NEC Corporation) [6]. For each interval, a BNM [7] is conducted from the ECDL results, and then a DBNM is obtained by connecting with the 3 BNMs via the "frontal" node in each BNM.

3 Results

Figure 3 shows a DBNM obtained for the present MRPs. The DBNM consisted of three BNMs, shoes intervals correspond to the E-BP, NS' and MP, respectively. Both of the BNMs for the E-BP and NS' intervals, contained the pre-supplementary motor area (pre-SMA) as the "frontal" node, and the SMA, the premotor cortex and the primary motor cortex as the "motor areas" one, while that for the MP one had the bilateral "somatosensory cortices" nodes which directly link from the frontal node, in addition to the "frontal" and "motor areas" nodes. These findings were strongly supported by the recent findings based on the scalp and epicortical recordings [8].

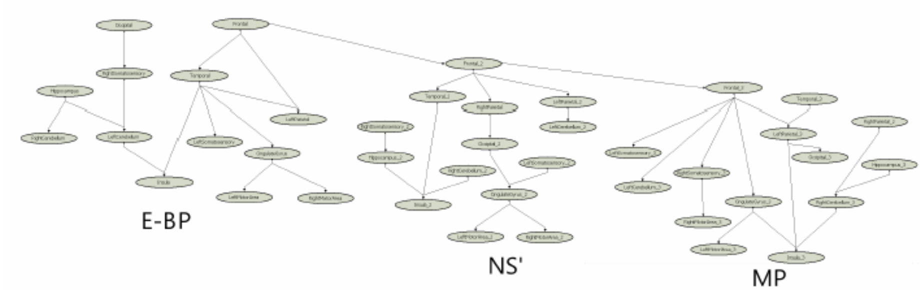


Fig. 3. A DBNM for the early-BP, the NS' and the MP

Our DBNM was obtained by directing the link from the frontal node in the E-BP-BNM to that in the NS'-BNM and the link from that in the NS'-BNM to that in the MP-BNM. In order to discriminate between the left- and right-hand movements from the single-trial EEGs recorded during the present tasks, the "left- and right- motor areas" *node activities* in the DBNM were calculated by the probabilistic inference [9] for each trial, and a difference in *node activities* between the nodes was examined by the t-test. This procedure was executed also for each of the E-BP-BNM, the NS'-BNM and the MP-BNM.

The results are summarized Table 1. The DBNM revealed the significant contralateral ERD (event-related desynchronization) in the right-hand movement ($p<.02$) and the contralateral ERS (event-related synchronization) in the left-hand movement ($p<.005$) as well as the MP-BNM ($p<.002$ and $p<.05$), while the E-BP-BNM ($p>.2$ and $p>.05$) and the NS'-BNM ($p>.9$ and $p>.2$) did not.

Table 1. t-tests for each BN and DBN

Interval	E-BP+NS'+MP				E-BP			
Movement	Right hand		Left hand		Right hand		Left hand	
Node	R-M1	L-M1	R-M1	L-M1	R-M1	L-M1	R-M1	L-M1
Mean	0.4374	0.2185	0.4271	0.1774	0.3682	0.1964	0.3027	0.1027
p-value	0.01177		0.004791		0.294		0.06507	

Interval	NS'				MP			
Movement	Right hand		Left hand		Right hand		Left hand	
Node	R-M1	L-M1	R-M1	L-M1	R-M1	L-M1	R-M1	L-M1
Mean	0.3684	0.3527	0.4457	0.2680	0.5755	0.1064	0.5330	0.1614
p-value	0.917		0.297		0.001072		0.04081	

4 Conclusion

The MP-BNM is different from the other BNMs in terms of network structures. This might suggest that there is a difference in the neural network in the brain between the movement onset and offset. By the t-test, it was proved to be difficult to discriminate between the left- and right-hand movement using only the E-BP-BNM and the NS'-BNM.

For the right-hand movement, the P-value in the t-test was minimal in the MP-BNM, compared with the other BNMs and the DBNM. However, the present DBNM could be totally more powerful for discriminating between the left- and right-hand movements, than the other BNMs including the MP-BNM.

Acknowledgement. This research was partly supported by a Grants-in-Aid for Scientific Research on Scientific Research (B) (20300196) – The Japan Society for the Promotion of Science.

References

1. Wolpaw, J., Birbaumer, N., McFarland, D., Pfurtscheller, G., Vaughan, T.: Braincomputer interfaces for communication and control. Clinical Neurophysiology 113, 767–791 (2002)
2. Sakamoto, M., Yamazaki, T.: A Bayesian network approach to single-trial-EEG-based BCI. In: Brain Topography and Multimodal Imaging, pp. 171–173. Kyoto University Press (2009)
3. Yamazaki, T., Sakamoto, M., Kamijo, K., Yamanoi, T.: A Bayesian network approach to brain-computer interface using independent component analysis and equivalent cament dipole source loealization. Neuroscience 2010 (40th Annual Meeting), Society for Neuroscience Final Program Tuesday, 79 (November 16, 2010)

4. Kamijo, K., Yamazaki, T.: A New Communication Channel Based on Brain Activity, Technical Report of IEICE HCS 1996-1997, pp.39–46 (1996)
5. Cichocki, A., Amari, S., Phan, A.H., et al.: ICALAB version 3 Toolbox. RIKEN BSI, http://www.bsp.brain.riken.jp/ICALAB
6. Kamijo, K., Kiyuna, T., Takaki, Y., Kenmochi, A., Tanigawa, T., Yamazaki, T.: Integrated approach of an artificial neural network and numerical analysis to multiple equivalent corrent dipole source localization. Frontier Med. Biol. Engng. 10(4), 285–301 (2001)
7. Cheng, J., Greiner, R., Kelly, J., Bell, D., Liu, W.: Learning Bayesian networks from data: an information-theory based approach. Artificial Intelligence 137(1-2), 43–90 (2002)
8. Shibasaki, H., Hallett, M.: What is the Bereitschaftspotential? Clinical Neurophysiology 117, 2341–2356 (2006)
9. Kabie, C.M., Hovel, D., Horvitz, E.: MSBNx: A Component-Centric Toolkit for Modeling and Inference with Bayesian Networks. Technical Report MSR-TR-2001-67 (2001)

Emotional Human-Machine Interaction: Cues from Facial Expressions

Tessa-Karina Tews, Michael Oehl, Felix W. Siebert,
Rainer Höger, and Helmut Faasch

Leuphana University of Lüneburg, Institute of Experimental Industrial Psychology,
Wilschenbrucher Weg 84a, 21335 Lüneburg, Germany
{tews,oehl,fsiebert,hoeger,pfister}@uni.leuphana.de

Abstract. Emotion detection provides a promising basis for designing future-oriented human centered design of Human-Machine Interfaces. Affective Computing can facilitate human-machine communication. Such adaptive advanced driver assistance systems (ADAS) which are dependent on the emotional state of the driver can be applied in cars. In contrast to the majority of former studies that only used static recognition methods, we investigated a new dynamic approach for detecting emotions in facial expressions in an artificial setting and in a driving context. By analyzing the changes of an area defined by a number of dots that were arranged on participants' faces, variables were extracted to classify the participants' emotions according to the Facial Action Coding System. The results of our novel way to categorize emotions lead to a discussion on additional applications and limitations that frames an attempted approach of emotion detection in cars. Implications for further research and applications are outlined.

Keywords: Emotion detection, human-computer interaction, human-centered design, affective computing.

1 Introduction

Although it is common knowledge that people are good at recognizing facial expression, it is not at all clear how algorithms can learn to encode or decode a human face in a real environment. Faces are complex, multidimensional, and they differ from each other. It is, therefore, a tremendous task to construct algorithms that enable computers to distinguish emotions just as well as humans can. In order to recognize facial expressions, we tried to implement a new dynamic approach in a realistic setting: in cars. The intention when driving is usually to reach a desired destination. En route, several events can cause the driver to feel intense emotions. Drivers can be annoyed by other road users, be nervous due to complex situations or just feel happy because all traffic lights seeming to be green. The effects on driving performance and accident risk caused by other influences such as alcohol, fatigue, smoking, and use of mobile phones during driving have been shown in several studies, e.g., [1] and [2]. The influence of emotions however has not been researched comprehensively. The

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 641–650, 2011.

importance of this topic can be derived from several statistics: every year about 39,000 people die in traffic accidents across Europe [3]. One approach to reduce the number of accidents is to detect mistakes made while driving and their real causes. During the more recent years emotions have drawn more and more attention, but to what extent emotions are responsible for accidents is still unresolved. However, Mesken and her colleagues reported that participants who reported feeling anger while driving tended to drive faster and exceeded the speed limit more often than participants who did not report similar feelings [4]. Nesbit et al. indicate that anger is associated with aggressive driving [5]. Most studies point out that happiness [4,6] and anger [5,7] are the most common emotions to influence how a car is driven [8].

For this reason this paper describes how neutral, happiness, and anger in facial expressions can be detected, and which methods can be applied to extract several emotions that correlate with certain facial expressions. We tried to influence the emotions within a driving context with up-to-date evaluated clips taken from movie sets in a real environment. After the driving contexts, we recorded a second set of artificial data. Here the participants were asked to facially express happy, angry and no emotions. Most of the current studies only apply static recognition methods that analyze the individual frames of a video. Furthermore, these experiments were mainly carried out in artificial settings, or were merely focused on anger while driving. Only very few studies cover multiple aspects, e.g., threatening driving [9] or personal violations [10]. To get an overview of current studies in this context Zeng et al. [11] presented a survey of audio, visual, and spontaneous affect recognition methods.

2 Method

In contrast to the majority of former studies that only make use of static recognition methods in rather artificial settings [12,13,14], we researched a new dynamic approach to detect emotions in facial expressions in an artificial setting and directly in a driving context. Within this context this paper tries to demonstrate how dynamic changes in neutral, happy and angry facial expressions can be analyzed. To prove this approach, an experiment was conducted with $N = 59$ (40 female) subject's at the age of M = 23.39 years (SD = 4.51). The facial expression data were recorded with a Logitech Quickcam Sphere camera (640x480 pixel) in a driving simulator STISIM DRIVE™ 2.0 as shown in figure 1. The videos were progressed for further analyses by independent raters in regard to the emotional content of the facial expressions. Variations were resolved by discussing the observed differences.

Fig. 1. The driving simulator

In order to record the driving videos, the participants had to drive along an ordinary road in the simulator as shown figure 2. Afterwards, a random movie was shown to each participant in order to trigger anger, happiness, or no certain emotion. This can be a possible efficient method to activate defined target emotions [15]. We used up-to-date evaluated clips out of movie sets which specified on and ranked the highest on the intended target emotion. The stimulus material that was used were clips taken from following movies: 'Hannah and her Sisters' (neutral), 'When Harry met Sally' (happiness / amusement) [both from 16] as well as 'Schindler's List' (anger) [17]. We evaluated the induction of emotions in pre-tests. Additionally, we applied the Self-Assessment Manikin (SAM) as a non-verbal pictorial assessment technique to measure participants' affective reactions [18,19]. The Self-Assessment Manikin (SAM) is of interest to distinguish between positive and negative emotional states.

After driving, we recorded a second set of artificial data where the participants were asked to show happy, angry and no facial expressions for 3 seconds in two trials.

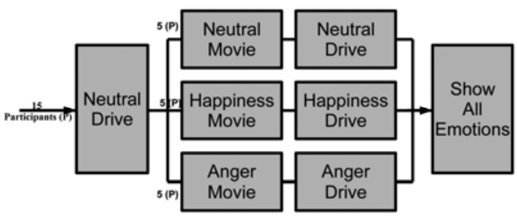


Fig. 2. Experimental procedure

During the entire experiment the participants were videotaped. To detect the facial expressions in this experiment, small dots were placed in their faces as shown in [20] and [21], but with fewer dots than in these studies. The dots (diameter 8 mm) were placed on 10 relevant facial muscles, e.g., figure 3 (a).

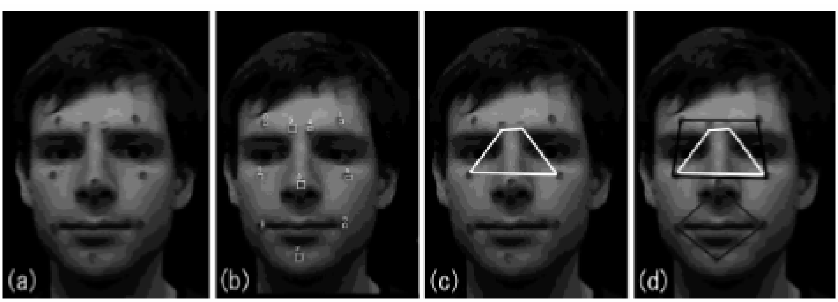


Fig. 3. **(a)** The positions of 10 relevant facial muscles, **(b)** Cluster detections, **(c)** The area defined by four blue dots, and **(d)** Three defined areas: the bigger above (area 1), the centered above (area 2), and the mouth area (area 3)

The exact positions of the dots were derived from earlier studies which had investigated facial muscle movement with the help of EMG, as well as observations of human mimics [22]. Since blue has the smallest part in the color-spectrum of the human face, blue was chosen as the color of the dots for the experiment. A software algorithm was developed which is able to detect the blue dots placed on the subjects' face by comparing special indicators of the recorded videos. The software detects the blue dots by searching each frame line by line for the blue color-spectrum and gathering all neighboring pixels to a cluster, see figure 3(b). By calculating the area between the dots, specific graphs of emotion can be generated and compared to each other. The variance of three areas between four dots was calculated dynamically over time, using an algorithm, see figure 3(c). 5 participants' videos were used to initialize the algorithm and the same number of videos were needed to test the software. In two major steps, we initialized the facial expressions' software: first, the algorithm was entered with the 5 participants' neutral, happy, and angry face data to distinguish between each emotion. The algorithm calculated the arithmetic mean of each emotion of the three areas as shown figure 3(d). Afterward the variance of each area was calculated using the arithmetic mean (see for example the neutral facial expression data formula 1):

$$Variance_{neutral}(x) = \frac{1}{N}\sum_{i=0}^{N}(mean_{neutral} - x_i)^2 \tag{1}$$

x_i = area size data

The algorithm summed up these initialized data and calculated the arithmetic mean of the area size with their variance for each emotion. Second, it calculated the three areas of the 5 remaining test participants' videos with their variance. During the next step the algorithm classified the emotions by comparing the test data video with the initialized data.

$$Index_{neutral}(y) = \frac{1}{N}\frac{\sum_{i=0}^{N}(mean_{neutral} - y_i)^2}{variance_{neutral}} \tag{2}$$

y_i = test area size data

This index is used to determine the variation from each test area size y_i to each emotion (see the second formula). The emotion with the smallest index value is the observed emotion. Finally, the results were presented in a graph.

3 Results

The Self-Assessment Manikin (SAM) is a non-verbal pictorial assessment technique that measures participants' affective reactions. Especially the valence of affective reactions is of interest to distinguish between positive and negative emotional states. For the first drive, i.e., the neutral route, participants across all experimental conditions reported no significant differences for the valence of emotional states ($F(2,55) = 2.06$; $p > .05$). During the second emotional route after the target emotion was induced by showing movie clips we observed a significant difference of the reported valence ($F(2,54) = 3.37$; $p < .05$). Compared to the neutral route the ratings for the emotional route in the neutral condition were on a ten-points-scale $M = 6.55$ (+0.00; $SD = 1.38$), in the happiness condition $M = 6.09$ (+0.25; $SD = 1.13$), and in the anger condition $M = 5.41$ (-0.59; $SD = 6.00$). This indicates by trend an effective emotion induction and a carry-over effect of the target emotion in the films for the emotional routes. Participants, however, showed no visible emotional expressions on the videotaped data and the driving data could not be used for our software evaluation.

The results in this paragraph were calculated from the second data set. Due to technical dropouts only $n = 10$ subjects' data could be analyzed. The graphs in figure 4 (neutral face video) and figure 5 (happy) show the histogram of a neutral and a happy face video. The histogram of an angry face video is very similar to the happy face video. Therefore, this histogram was not included.

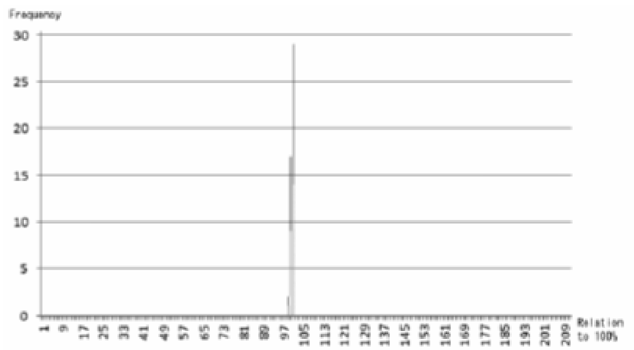


Fig. 4. A neutral face video results

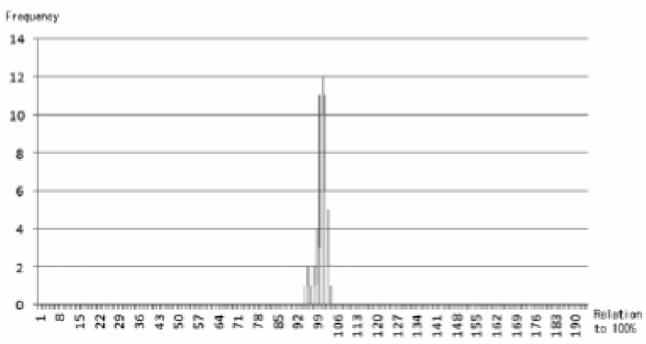


Fig. 5. A happy face video results

The arithmetic mean of each area is set to 100% and each value was summed up. The frequency was summed up on the x-axis and describes how often an area has this specific size. The average relation is shown on the y-axis and reflects the size of an area. The raw data of the neutral face video does not show any strong variation of the arithmetic mean. All values are concentrated around 100%, because there have been few movements in the face as expected in a neutral face. In contrast, the happy and angry face videos show stronger facial movements, which the program can use to distinguish between facial expressions. The areas were extensively changing in size because of the mimic activities caused by the emotions. Angry areas seem to vary a little bit more than the happy faces do.

In our experiment 10 participants' faces were recorded. The evaluation of the videos revealed that participants did not show any visible facial expressions of emotions during the driving situations. Therefore, we could only evaluate the second data setting of the artificial emotions instead of the driving videos. The graphs in figure 6 show the average of the three area sizes of the participants' neutral, happy, and angry facial expressions. The participants' average area size is displayed on the x-axis and the participants number is shown on the y-axis. Observing the size of the upper areas 1 and 2 the values were quite similar and our software was not able to distinguish between each emotion. Therefore, we could only analyse the mouth area. Each emotion's arithmetic mean of area size was neutral facial expressions (5555 pixel), happy facial expressions (6730 pixel), and angry facial expressions (4704 pixel). This relatively large gap helped distinguish the emotions happy and angry. The non specific facial expression were classified with the low variance values as shown figure 7.

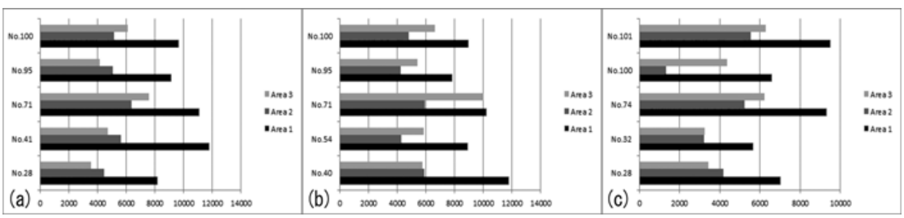


Fig. 6. The initialize data average of the area size with **(a)** neutral, **(b)** happy, and **(c)** angry facial expression

The graphs in figure 9 show the average of the three area variances of the participants' neutral, happy, and angry facial expressions. The participants' average variance size is displayed on the x-axis and the participants number is shown on the y-axis. As expected, the data of the non specific facial expression showed the lowest variance. This result helped to easily discriminate the non specific facial expression videos from the other emotions. The variance of happy and angry facial expressions did not help to distinguish these emotions.

The remaining 5 artificial faces' videos were compared to these initialized data sets for each emotion (happy, angry and neutral). Figure 8 shows the average of the mouth area variance of the test participants' neutral, happy, and angry facial expressions. The participants' average variance size is displayed on the x-axis and the participants themselves are shown on the y-axis. By comparing the initial videos to the test videos happy and angry facial expressions could be discriminated.

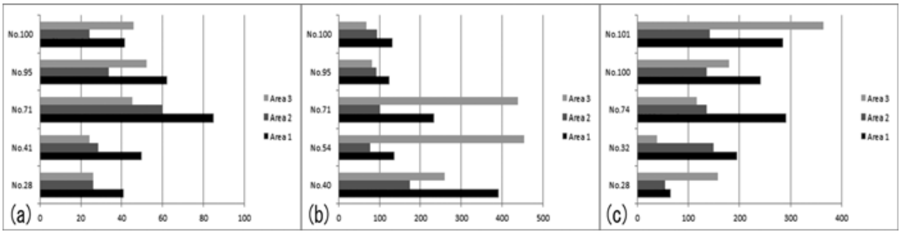


Fig. 7. The initialize data average of the variance with **(a)** neutral, **(b)** happy, and **(c)** angry facial expression

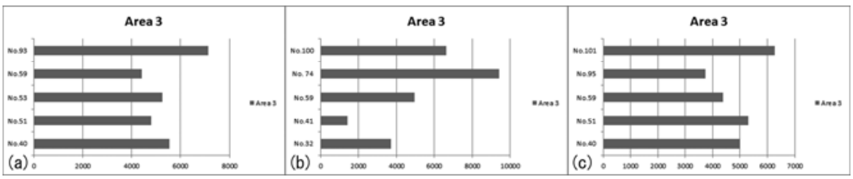


Fig. 8. The test data average of the area size with **(a)** neutral, **(b)** happy, and **(c)** angry facial expression

The graphs in figure 9 displays the average variance of the mouth area calculated from the participants' neutral, happy, and angry facial expressions. The participants' average area variance is shown on the x-axis and the participants are displayed on the y-axis. Again the data of the non specific facial expression showed the lowest variance. For this reason no facial expression videos could be discriminated from the other emotions.

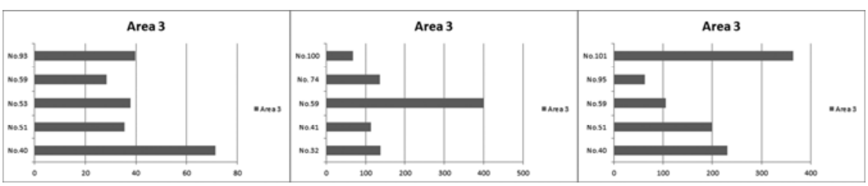


Fig. 9. The test data average of the variance with **(a)** neutral, **(b)** happy, and **(c)** angry facial expression

By analyzing the change of the three areas, the mouth area (area 3) yielded to the best results. By observing changes of the mouth areas we enabled our software to distinguish between angry and happy faces. Neutral faces could be detected by the low variance values of the videos. The other two areas in the upper part of the face did not contribute substantially to the classification of the emotions.

4 Discussion

As reviewed from other literature, the algorithms for facial expressions are normally quite complex [11]. In contrast to this approach, we tried to minimize complex computations by using a simple feature extraction method. Our software presented in this paper calculates the variance of three areas between four dots in the face over time as a new dynamic approach for detection of emotions by facial expressions, using an autonomous algorithm. By computing the facial expression data of the mouth area with the index of emotions and the variance of this areas over time, our software can classify five participant's emotions in happy, angry and neutral if the emotions are expressed strong enough. In our experiment n = 10 participants' faces were analyzed. 5 videos were used to calculate the arithmetic mean and the average variance of each emotion. The evaluation of the remaining videos showed that the more obvious the motions are and more sources of motion can be found in the image, the more sources of discrimination between emotions can be detected. Additionally, participants showed little emotion to any of the road situations. Those participants' videos were not included in the results of this paper. The main drawbacks during our research were the small amounts of comparing and testing data and that our software was very sensitive to any variance of head orientations. A possible explanation for the limited expression of emotions is that the participants felt very safe inside the driving simulator and that they have been blunted by the movies. We evaluated the induction of emotions in pre-tests which showed that the induction was successful. However, participants showed no visible emotional expressions on the videotaped data as shown in other driving simulator studies [4]. The videos were progressed for further analysis by independent raters in order to find emotions within the participants' facial expression. If the people analyzing the videos talk or yawn could influences the analysis of facial expression. However, in further studies the multi-modal approach with speech analysis or other methods, e.g., the Mahalanobis-Distance [23], can help to distinguish these facial expressions.

Similar to other facial expression studies, we had problems with hair or head orientations [24, 25]. Additionally, screen reflections temporarily covered the blue dots and the software was not able to calculate the area correctly. In order to reduce the hair problem, the participants were asked to wear an alice band. The participant's facial expression might also have been inhibited by the blue points. To avoid this, the capabilities of the facial feature extraction could be enhanced in the future, making the blue points obsolete.

In our experiment the conditions were constructed artificially by having the setting in a laboratory environment. In a real driving situation the variance of lighting conditions and reflections could inhibit the classification the drivers' emotions. Future steps could be the detection of emotions with a broader range of non-invasive measures, e.g., speech recognition, heart rate, and grip-strength will be pursued. To follow our future centered multi-modal approach, simultaneous speech data was recorded in our experiment which will be analyzed in the next step. To enhance the generalization of the results, the software will be tested on a larger database with a lot more subjects and in real driving situations.

Acknowledgments. This research was funded by grant FS-Nr. 2006.63 from the 'Arbeitsgruppe Innovative Projekte' (AGIP) of the Ministry of Science and Culture, Lower Saxony, Germany. Many thanks go to Patricia Brunkow, Johanna von Stebut, Raphael Seitz, Christoph Ruth and Marco Wiethof for their help in conducting the experiment as well as to Heiko Witthöft for his help in analyzing the experiment and his support in applied computer science.

References

1. Crancer, A., Dille, J.M., Delay, J.C., Wallace, J.E., Haykins, M.D.: Comparison of the effects of marijuana and alcohol on simulated driving performance. Science 164, 851–854 (1969)
2. Weiler J.M., Bloomfield J.R., Woodworth G.G., Grant, A.R., Layton, T.A., Brown, T.L., McKenzie, D.R., Baker, T.W., Watson, G.S.: Effects of fexofenadine, diphenhydramine, and alcohol on driving performance. A randomized, placebo-controlled trial in the Iowa driving simulator. Annals of Internal Medicine 132, 354–363 (2000).
3. dpa Deutsche Presse-Agentur: Weniger Verkehrstote auf Europas Straßen. Frankfurter Rundschau, `http://www.fr-online.de/auto/weniger-verkehrstote-auf-europas-strassen/-/1472790/3076970/-/index.html` (retrieved June 23, 2009)
4. Mesken, J., Hagenzieker, M.P., Rothengatter, T., De Waard, D.: Frequency, determinants, and consequences of different drivers' emotions: An on-the-road study using self-reports (observed) behaviour, and physiology. Transportation Research Part F 10, 458–475 (2007)
5. Nesbit, S.M., Conger, J.C., Conger, A.J.: A quantitative review of the relationship between anger and aggressive driving. Aggression and Violent Behavior 12(2), 156–176 (2007)
6. Levelt, P.B.M.: Praktijkstudie naar emoties in het verkeer (Emotions in Traffic). SWOV Report R-2003-08. SWOV, Leidschendam, Netherlands (2003)
7. Deffenbacher, J.L., Oetting, E.R., Lynch, R.S.: Development of a driving anger scale. Psychological Reports 74, 83–91 (1994)
8. Oehl, M., Roidl, E., Frehse, B., Suhr, J., Siebert, F.W., Pfister, H.-R., Höger, R.: Das Emotionsspektrum von Autofahrern im Straßenverkehr. In: Petermann, F., Koglin, U. (eds.) (Hrsg.), 47. Kongress der Deutschen Gesellschaft für Psychologie, Abstracts, pp. 371–372. Pabst Science Publishers, Lengerich (2010)
9. Wells-Parker, E., Ceminsky, J., Hallberg, V., Snow, R.W., Dunaway, G., Guiling, S., Williams, M., Anderson, B.: An exploratory study of the relationship between road rage and crash experience in a representative sample of US drivers. Accident Analysis and Prevention 34, 271–278 (2002)
10. Mesken, J., Lajunen, T., Summala, H.: Interpersonal violations, speeding violations and their relation to accident involvement in Finland. Ergonomics 45, 469–483 (2002)
11. Zeng, Z., Pantic, M., Roisman, G.I., Huang, T.S.: A Survey of Affect Recognition Methods: Audio, Visual, and Spontaneous Expressions. IEEE Transactions on Pattern Analysis and Machine Intelligence 31(1), 39–58 (2009)
12. Kirby, M., Sirovich, L.: Application of the Karhunen-Loeve procedure for the characterization of human faces. IEEE Transactions on Pattern Analysis and Machine Intelligence 12, 103–108 (1990)
13. Sirovich, L., Kirby, M.: Low-dimensional procedure for the characterization of human faces. Journal of the Optical Society of America A 4, 519–524 (1987)
14. Lin, C.-H., Wu, J.-L.: Automatic Facial Feature Extraction by Genetic Algorithms. IEEE Transactions on Image Processing 8, 834–845 (1999)

15. Rottenberg, J., Ray, R.D., Gross, J.J.: Emotion Elicitation Using Films. In: Coan, J.A., Allen, J.J.B. (eds.) Handbook of Emotion Elicitation and Assessment, pp. 9–28. Oxford University Press, Oxford (2007)
16. Hewig, J., Hagemann, D., Seifert, J., Gollwitzer, M., Naumann, E., Bartussek, D.: A revised film set for the induction of basic emotions. Cognition & Emotion 19/7, 1095–1109 (2005)
17. Schaefer, A., Nils, F., Sanchez, X., Philippot, P.: Assessing the effectiveness of a large database of emotion-eliciting films: A new tool for emotion researchers. Cognition & Emotion 24/7, 1153–1172 (2010)
18. Lang, P.J.: Behavioral treatment and bio-behavioral assessment: computer applications. In: Sidowski, J.B., Johnson, J.H., Williams, T.A. (eds.) Technology in Mental Health Care Delivery Systems, pp. 119–137. Ablex, Norwood (1980)
19. Bradley, M.M., Lang, P.J.: Measuring Emotion: The Self-Assessment Manikin and the Semantic Differential. Ther. & Exp. Psychiat. 25/1, 49–59 (1994)
20. Kaiser, S., Wehrle, T.: Automated coding of facial behavior in humancomputer interactions with FACS. Journal of Nonverbal Behavior 16, 67–83 (1992)
21. Bassili, J.N.: Emotion recognition: the role of facial movement and the relative importance of upper and lower areas of the face. Journal of Personality and Social Psychology 37, 2049–2058 (1979)
22. Cohn, J.F., Ekman, P.: Measuring facial action by manual coding, facial EMG, and automatic facial image analysis. In: Harrigan, J.A., Rosenthal, R., Scherer, K. (eds.) Handbook of Nonverbal Behavior Research Methods in the Affective Sciences, pp. 9–64. Oxford University Press, New York (2005)
23. Mahalanobis, P.C.: On the generalised distance in statistics. Proceedings of the National Institute of Science of India 12, 49–55 (1936)
24. Lin, C.-H., Wu, J.-L.: Automatic Facial Feature Extraction by Genetic Algorithms. IEEE Transactions on Image Processing 8, 834–845 (1999)
25. Agarwal, M., Jain, N., Kumar, M., Agrawal, H.: Face Recognition using Principle Component Analysis, Eigenface and Neural Network. In: Jha, M., Long, C., Mastorakis, N., Bulucea, C.A. (eds.) Sensors, Signals, Visualization, Imaging, Simulation And Materials, pp. 204–208. The World Scientific and Engineering Academy and Society, Wisconsin (2009)

Development of an Eye-Tracking Pen Display for Analyzing Embodied Interaction

Michiya Yamamoto[1], Hiroshi Sato[1], Keisuke Yoshida[1],
Takashi Nagamatsu[2], and Tomio Watanabe[3]

[1] 2-1 Gakuen Sanda Hyogo 669-1137, Japan
{michiya.yamamoto,hiroshi.sato,keisuke.yoshida}@kwansei.ac.jp
[2] 5-1-1 Fukae-minami Higashi-nada Kobe Hyogo 658-0022, Japan
nagamatu@kobe-u.ac.jp
[3] 111 Kuboki Soja Okayama 719-1197, Japan
watanabe@cse.oka-pu.ac.jp

Abstract. In recent times, intuitive user interfaces such as the touch panel and pen display have become widely used in PCs and PDAs. Previously, the authors developed the *bright pupil camera*. They subsequently developed an eye-tracking pen display based on this camera and a new aspherical model of the eye. In this paper, a robust gaze estimation method that uses a integrated-light-source camera is proposed for analyzing embodied interaction. Then, a prototype of the eye-tracking pen display was developed. The accuracy of the system was approximately 12 mm on a 15" pen display, which is sufficient for human interaction support.

Keywords: Embodied interaction, pen display, eye-tracking, aspherical model.

1 Introduction

Today, the intuitive user interfaces of PCs and PDAs, such as touch panel and pen display, have become widely used. These devices are expected to open up a new embodied interaction and communication as well as interaction between humans and computers.

By focusing on the importance of embodied interaction, the authors have developed a Computer Graphics (CG)-embodied communication support system [1]. Especially, the importance of timing control in generating embodied motions and actions is made clear for supporting natural, familiar, and polite interaction via CG and robot agent [2]. However, for making further uses of embodiment, it is required to analyze the relationships between body motion and attention.

If we could integrate pen display and eye-tracker, it becomes possible to analyze various embodied interactions. For example, we could analyze how a presenter indicates or emphasizes a slide in presentation by using intuitive pen display. In addition, such an eye-tracking pen display could become a gadget for realizing a new mode of interaction between humans and computers.

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 651–658, 2011.

The authors have already developed a prototype of Eye-Tracking Pen Display, MobiGaze which enables eye-tracking on iPhone, and ETTI (Eye-Tracking Tabletop Interface) [3][4][5]. The accuracy was about 0.7°, which was enough for interaction analysis, however, the robustness of the system was not enough. In this study, we have developed a robust pupil detection method by using aspherical model of the eye and dark pupil method. The evaluation experiment shows the effectiveness of new prototype system.

2 Technical Requirements

There are several eye-trackers which we can be listed as de facto standards such as Tobii T120 [6]. However, they are not suitable for use with pen display. The biggest problem of such eye-trackers is that they have cameras and IR LEDs under their displays (Fig. 1). When a right handed person use a pen on the display, the right arm may hide the camera or LED.

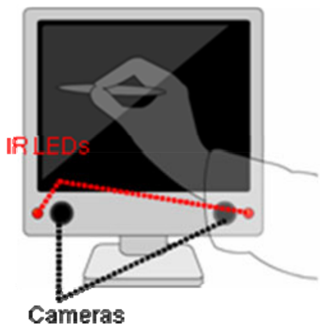


Fig. 1. Typical layout of cameras and LEDs

The tracking distance and gaze angle may also cause a problem when a user draws on a pen display. Because, the tracking distance of existing eye-trackers is approximately 50 cm or more, and the gaze angle is approximately 30° in many cases. If we put an eye-tracker at the left bottom of the display and use a pen on the display, the tracking distance becomes too close and gaze angle becomes too wide.

In addition, easy calibration is required for eye-tracking pen display, so that intuitive interface can be realized. Thus, we can summarize the technical requirements as follows [3]:

- Free arrangement of cameras and LEDs to prevent obstruction by the right hand
- Robust gaze estimation with short distance & wide gaze angle
- Easy calibration

3 Arrangement of Cameras and IR LEDs

We reviewed previous studies and developed a prototype of the system by considering its technical requirements. The 3D gaze-tracking approach was selected for accuracy [7][8][9]. This approach involves the use of two cameras and three or four LEDs. Fig. 2 (a) shows the arrangement of the system proposed by Nagamatsu et al [10]. In this study, we first developed a prototype of the system by positioning the cameras and LEDs: two cameras are placed to the left of the pen display, and one LED each is placed on the top, left, and bottom frames of the pen display (Fig. 2 (b)). However, even with such an arrangement, stable eye-tracking cannot be realized due to the obstructions by the right hand and the eyelid. Therefore, we reviewed the arrangements proposed in previous studies again. Some researchers have proposed camera-LED integrated systems. For example, Ohno developed a system that involved the use of one camera and two LEDs [11]. Chen et al. developed a system that involved the use of two cameras and two LEDs mounted near the camera centers; in this arrangement, the camera and the LED were integrated into one component [12]. We can arrange such a system to the left of the pen display (Fig. 2 (c)); however, such a system would be inadequate if the pen display is to be used at various angles. The two cameras should be separated for the eye tracking pen display system [3].

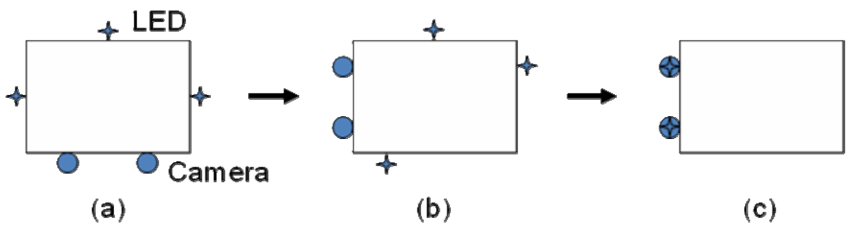


Fig. 2. Arrangement of cameras and LEDs

4 Estimation Method of the Optical Axis of the Eye

Here, we estimated the optical axis by using two cameras and two IR LEDs by using the arrangement as shown in chapter 3. First, we calibrated the external parameters of the cameras to acquire the relative 3D position of the cameras and the displays. Next, on the basis of the results of image processing, we detected the position of the pupil and two bright points as a Purkinje image as shown in Fig. 4 (a). We assumed that the light source and the camera center were at the same position. Then, we obtained a plane that contained vectors $\mathbf{A}$ and $\mathbf{B}$, as shown in Fig. 3 by using the expression $(\mathbf{C}_j - \mathbf{B'}_j) \times \mathbf{P'}_j - \mathbf{C}_j) \cdot (\mathbf{X} - \mathbf{C}_j) = 0$, where $\mathbf{X}$ is a point on the plane. One camera and one LED were used for determining a plane that contained the optical axis. Therefore, the optical axis could be obtained as the intersection of two planes by using the two cameras and two LEDs. Then, the user gazed at a point on the tabletop interface for the calibration. The difference between the optical axis and the visual axis of the eye was revised by carrying out this calibration. The cross point of the visual axis and the pen tabletop was estimated as the gaze point [3].

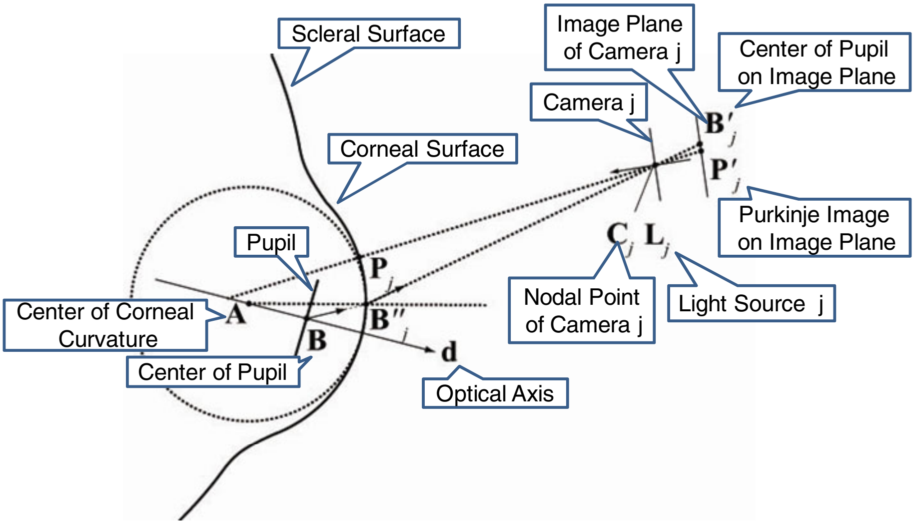


Fig. 3. Estimation of the optical axes

5 Image Processing

We have developed an eye-tracking pen display and achieved eye tracking of up to approximately 60° on the basis of our estimation method of the optical axis of the eye, which is introduced in chapter 4, using a robust layout of cameras and LEDs [3]. However, because of the unstable estimation of the optical axis, which is caused by image processing error in bright pupil method, as shown in Fig. 4 (b), the analysis of embodied interaction was difficult. Although the dark pupil technique with one IR LED around lens proved reliable (Fig. 4 (c)), it is necessary to develop a new pupil detection method because the distance between the LED and camera center causes error.

(a) Bright Pupil (b) Image Processing Error (c) Dark Pupil

Fig. 4. Image processing

6 Pupil Detection Method

In this study, the authors developed integrated-light-source cameras that had IR LEDs around their lenses. By calculating the barycentric position of two or three of the LEDs around the lens, we simulated a virtual light source (Fig. 5). These IR LEDs were controlled by Arduino microcontroller [13]. To evaluate the effectiveness of our technique, we performed an experiment wherein we compared the results of pupil detection technique using light integrated cameras with one, two, and three light sources and those of the bright pupil technique. Fig. 6 shows the system configuration.

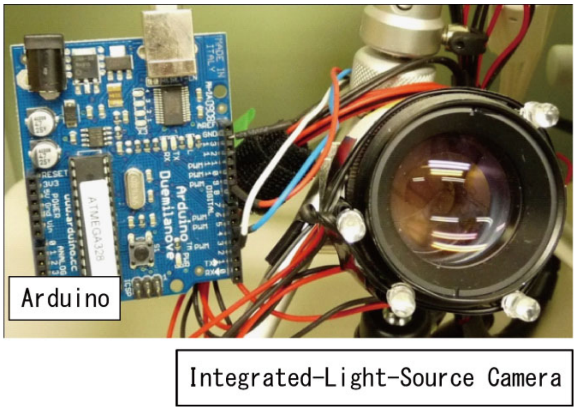


Fig. 5. Integrated-light-source camera

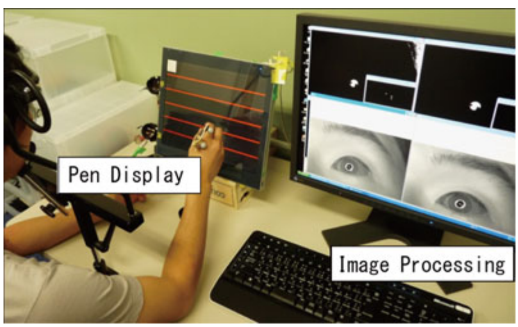


Fig. 6. Prototype of eye-tacking pen display with integrated-light-source cameras

In the experiment, we lit the IR LEDs, positioned the subject's head on a chin rest, and asked the subject to look at the marker on the left side of the pen display for calibration. Next, we displayed a white cross on the pen display and asked the subject to gaze at the center of the white cross for 10 frames. The cross was displayed on each of the 128 pixels in the right side and each of the 192 pixels in the left side. The center of the cross was displayed more to the right because a large margin of error was expected for the light integrated cameras set up on the left side of the pen display. In this manner, we evaluated four patterns (one, two, three LEDs, and bright pupil). The same light volume was ensured for the LED patterns by adjusting the light volume using a phototransistor. Five unaided-eye students participated in the experiment.

Fig. 7 shows the results. There was no significant difference between the results of two LEDs, three LEDs, and bright pupil method, as suggested by one-way analysis of the variance to average error (Fig. 8). Therefore, we decided to use two LEDs. The processing speed of our system was 9.1 fps with a computer having a Intel Core 2 Duo 3.0 GHz processor, and the delay between human action and gaze measurement was 7/30 frame.

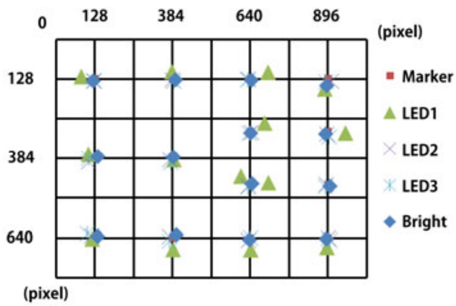


Fig. 7. Result of evaluation experiment

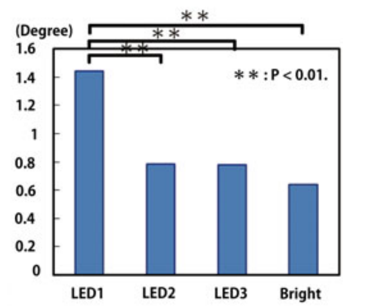


Fig. 8. Result of one-way analysis of variance

7 Assessment of Accuracy for Dynamic Gaze Tracking Accuracy

We also performed an experiment to analyze gaze interaction. Here, we asked the user to gaze at a moving marker. The speed of the moving marker was 284 pixel/s that human being drew a line. This speed was lower than that of the saccades. In addition, there were two kinds of marker move: one was the drawing of a pattern of five lines (of length from 128 pixel to 896 pixel) and the other was the drawing of the figure "Z." The average error for the five-line pattern was 42.9 pixel (12.4 mm) and that for the "Z" pattern was 48.1 pixel (14.3 mm). Fig. 9 shows an example of the results. Fig. 10 shows the interaction timing of the five-line patterns and the gaze points of the subjects. The x-axis shows the frame on following the marker, and the y-axis shows the pixel on the screen. There was a gaze position lag of approximately 0.66 s (average value for three subjects) against the movement of the marker. Fig. 11 shows the interaction timing of figure "Z." The average lag was approximately 0.61 s.

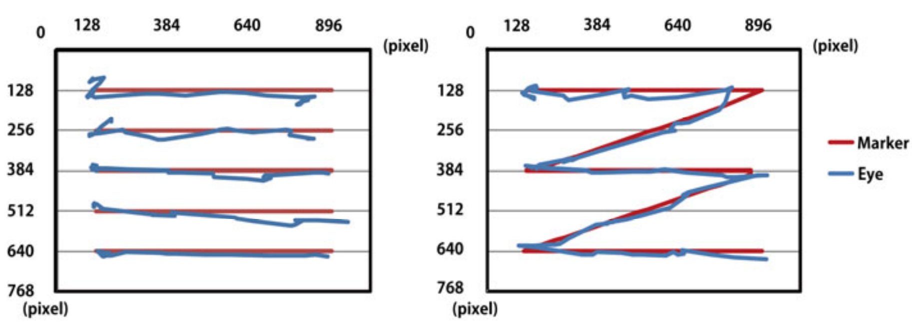


Fig. 9. Example of results

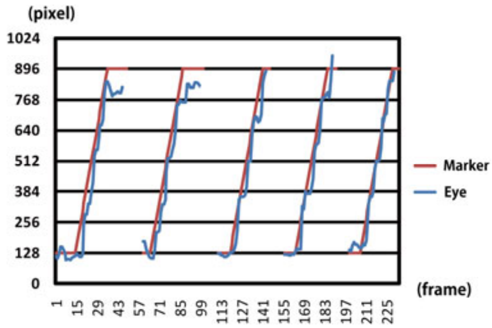


Fig. 10. Example of results of x-axis of 5-line pattern

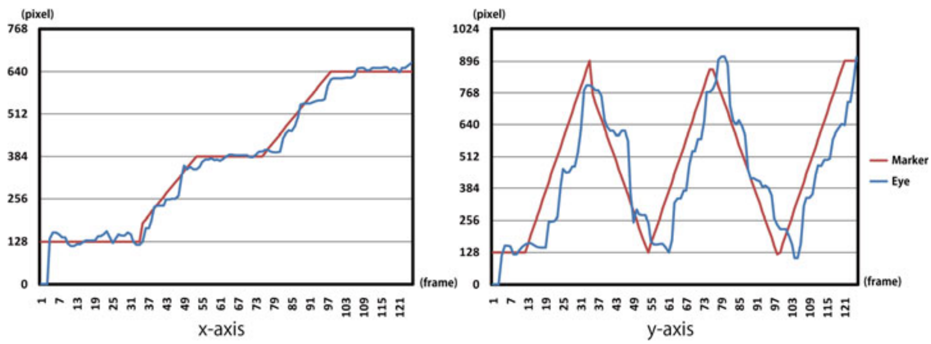


Fig. 11. Results of figure "Z" pattern

8 Conclusion

In this paper, a robust gaze estimation method was proposed for analyzing embodied interaction, and a prototype of an eye-tracking pen display based on integrated-light-source cameras was developed. The experimental results show that the accuracy of the system was approximately 12.4 mm for the 5-line pattern and 14.3 mm for the figure "Z" pattern on a 15" pen display. Because these were measured in an embodied interaction experiment, it is clear that the accuracy was sufficient for human interaction analysis.

Acknowledgements. This work under our project "Embodied Communication Interface for Mind Connection" has been supported by "New IT Infrastructure for the Information-explosion Era" of MEXT Grantin-Aid for Scientific Research on Priority Areas. Also, our project "Generation and Control Technology of Human-entrained Embodied Media" has been supported by CREST (Core Research for Evolution Science and Technology) of JST (Japan Science and Technology Agency).

References

1. Yamamoto, M., Watanabe, T.: Development of an Embodied Interaction System with InterActor by Speech and Hand Motion Input. In: CD-ROM of the 2005 IEEE International Workshop on Robots and Human Interactive Communication, pp. 323–328 (2005)
2. Yamamoto, M., Watanabe, T.: Timing Control Effects of Utterance to Communicative Actions on Embodied Interaction with a Robot and CG Character. International Journal of Human-Computer Interaction 24/1, 103–112 (2008)
3. Yamamoto, M., Nagamatsu, T., Watanabe, T.: Development of eye-tracking pen display based on stereo bright pupil technique. In: Proceedings of the 2010 ACM Symposium on Eye-Tracking Research & Applications, pp. 165–168 (2010)
4. Yamamoto, M., Komeda, M., Nagamatsu, T., Watanabe, T.: Development of Eye-Tracking Tabletop Interface for Media Art Work. In: Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces 2010, pp. 295–296 (2010)
5. Nagamatsu, T., Yamamoto, M., Sato, H.: MobiGaze: Development of a Gaze Interface for Handheld Mobile Devices. In: Proceedings of the 28th of the International Conference Extended Abstracts on Human Factors in Computing Systems, pp. 3349–3354 (2010)
6. Tobii, `http://tobii.com/`
7. Shih, S.-W., Liu, J.: A Novel Approach to 3-D Gaze Tracking using Stereo Cameras. IEEE Transactions on Systems, Man, and Cybernetics Part B 34/1, 234–245 (2004)
8. Guestrin, E.D., Eizenman, M.: Remote Point-of-Gaze Estimation with Free Head Movements Requiring a Single-Point Calibration. In: Proceedings of the 29th Annual International Conference of the IEEE EMBS, pp. 4556–4560 (2007)
9. Nagamatsu, T., Iwamoto, Y., Kamahara, J., Tanaka, N., Yamamoto, M.: Gaze estimation method based on an aspherical model of the cornea: surface of revolution about the optical axis of the eye. In: Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications, pp. 255–258 (2010)
10. Nagamatsu, T., Kamahara, J., Tanaka, N.: 3D Gaze Tracking with Easy Calibration Using stereo Cameras for Robot and Human Communication. In: Proceedings of IEEE ROMAN 2008, 59–64 (2008)
11. Ohno, T.: One-point calibration gaze tracking method. In: Proceedings of the 2006 Symposium on Eye Tracking Research & Applications, vol. 34 (2006)
12. Chen, J., Tong, Y., Gray, W., Ji, Q.: A Robust 3D Eye Gaze Tracking System using Noise Reduction. In: Proceedings of the 2008 Symposium on Eye Tracking Research & Applications, pp. 189–196 (2008)
13. Arduino, `http://www.arduino.cc/`

Care Giving System Based on Consciousness Recognition

Noriko Yokoyama, Tomoyuki Yamaguchi, and Shuji Hashimoto

Department of applied physics, Waseda University,
3-4-1 Okubo, Shinjuku-ku, Tokyo, Japan
yokoyama@shalab.phys.waseda.ac.jp

Abstract. In these days, robotics systems that provide supportive communication to human have been actively developed. However in such systems the internal consciousness state of human is not taken into consideration and hence the provision of support might not be appropriate. In this article we proposed a study support communication system that encourages and praises the human user based on the recognition of consciousness state through the user posture.

Keywords: Human interface, Bayesian network.

1 Introduction

The robots that support human life such as a humanoid robot and a pet robot have been actively developed. These robots have to observe human behavior to provide appropriate supports. In such a system, if the human internal state is recognized through the person's behavior, the robot will be able to offer more suitable support to the people. Moreover, while the development of the robot that supports people's daily life is advanced, it is pointed out that we do not feel like the robot blending into society [1]. To develop a more familiar robot, we should construct a relation between people and a robot [2][3] and communication between people and robots is important. The essence of communication is the exchange and sharing of state of mind [4]. It is necessary that we provide a robot with an internal state of human and a function that makes the robot seems like having mind and it is important how a robot extracts internal state of human from this point of view.

As one of the systems that observe people to support them, there is a cooperative system of the environmental type robot and the individual robot, where the location information of people in the house is detected by pressure sensors spread over the floor and infrared cameras. The location information is used for an appliance operation [5] or a physically support like taking something [6]. These systems can support people, but the internal state of the people is not considered. While, as the research that estimates the internal state of people, there is a sleep detection system while driving [7]. The consciousness level is estimated by detecting the frequency and opening time of blinks using images captured with the infrared camera for warning.

In this article, we developed a system that provides supportive communication based on consciousness recognition of the user. We show the details of the proposed system and the experimental results.

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 659–668, 2011.

2 Proposed System

The proposed system is composed of three parts as shown in Figure 1. First, the images of a user are taken with a USB camera and the posture of the user is estimated. Secondly, the user's state of consciousness is estimated from the time-sequence data of the posture using neural network. Finally, the system selects an utterance and utters to the user based on the time-sequence of the state of consciousness using Bayesian network.

2.1 Posture Estimation

We extract the silhouette of the user from an image using background difference method and estimate the posture using template matching. We prepared four types of postures "normal (upright) ", "leaning", "downward" and "face down ".

(1) Acquisition of Silhouette. User's silhouette is extracted by background difference method of gray scale image. The size of input image is 640×480 pixels. The library function of OpenCV is used for making grayscale image. The pixel value of (x, y) of pre-processed image is defined as f(x, y) and pixel value of (x, y) of post-processed image is defined as g(x, y) and we obtain silhouette image using Equation 1 where the threshold value is defined as T.

$$g(x,y)=\begin{cases}255 & f(x,y)>T\\ 0 & f(x,y)\le T\end{cases} \tag{1}$$

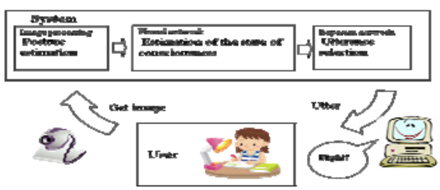


Fig. 1. Overview of system

1 image average of 100 images

Fig. 2. Example of typical image

(2) Generation of Template Image. Four kinds of typical images are generated by averaging each pixel value on g(x, y) of 100 images of each posture group. These average images are used in next template matching. Figure 2 shows the example of typical image and the template image of "normal" obtained by averaging.

(3) Template Matching. The distances between silhouettes of input image P(*x*, *y*) of the user and each template image $Q_i(x, y)$ are calculated by Equation 2. The template image that has the smallest distance is detected. Its posture is estimated as the posture of the user.

$$d = \frac{1}{N_W} \sum_{(x,y)\in W} \left| P(x,y) - Q_i(x,y) \right| , \quad (2)$$

where *i* denotes the posture number.

To prevent the influence of the input situation on the desk, the range of silhouette W is defined as 640×420 pixels. N_W is the number of pixels in area *W*.

2.2 Estimation of State of Consciousness

Above-mentioned posture estimation is carried out at 2fps and the state of consciousness is estimated based on the temporal feature of the posture changes. As an input, we converted five minutes of posture data (length 600) into ten parameters composed of the total number of each posture:4 and the number of posture transition:6. As an output, we prepared four levels of consciousness state, ”concentrating”, ”sleepy”, ”lack the drive” and ”sleeping”. These levels were selected based on phase theory [8].

(1) Phase Theory. Phase theory was introduced by Kunie Hashimoto in 1978. This theory divides the consciousness state into five levels from phase 0 to phase Ⅳ (Table 1). In this research, four consciousness state (Table 2) was prepared referring to above-mentioned phase 0 to Ⅲ.

(2) Neural network learning. The learning of each state of consciousness is performed with a three-layered perceptron with back propagation algorithm (Figure 3). In this research, neural network was chosen because the pattern of postures are classified into either of four states of consciousness. The reason we used posture data with time width is that consciousness state can be read from continuous actions. Time length is five minutes in this research because too long length cannot correspond to the change of the state of consciousness.

Table 1. Stage of consciousness state

Phase	Mode of consciousness	Physiology
0	unconscious, trance	sleeping, brain attack
I	Subnormal	fatigue, doze ,drunken
II	normal, relaxed	Relaxed
III	normal ,clear	Active
IV	hypernormal, excited	Panic

Table 2. Consciousness states

Phase theory		Consciousness states
0 : sleeping	→	「sleeping」
Ⅰ : doze	→	「sleepy」
Ⅱ : relaxed	→	「lack the drive」
Ⅲ : active	→	「concentrating」

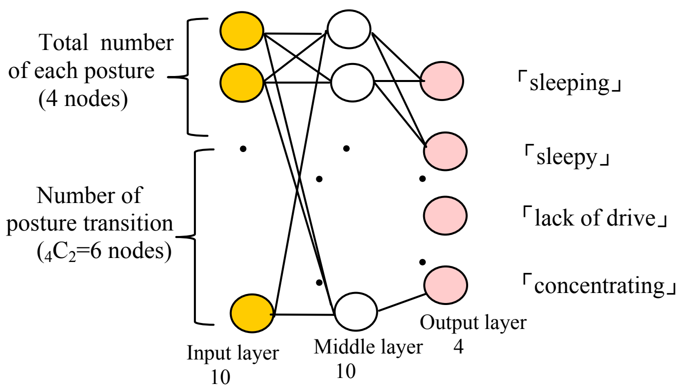


Fig. 3. Neural network

Table 3. Values of output

Consciousness states	Value of output
sleeping	(1,0,0,0)
sleepy	(0,1,0,0)
lack of drive	(0,0,1,0)
concentrating	(0,0,0,1)

As an input, we converted five minutes of posture data (length 600) into ten parameters composed of the total number of each posture and the number of posture transition. We use not only total number but also transition frequency between each posture, so time series information can be considered.

The teacher data is made by providing a state of five minutes of posture data and we let neural network learn the pattern. The output of each state is given as shown in Table 3. The learning of each state of consciousness is performed 10000 times. The average of error compared with learning data is

$$E = 3.734e^{-5} \fallingdotseq \quad 0.025. \tag{3}$$

(3) Estimation of state of consciousness. We input converted ten parameters into learned neural network and the highest output is the estimated state at the time.

2.3 Utterance Selection

The system selects an utterance according to the time-sequence of consciousness using Bayesian network. Figure 4 shows the model of Bayesian network. There are 9 situation nodes and 10 utterance nodes. Table 4 shows situation nodes. These are calculated based on the time-sequence consciousness. For example, if the user follows the following path, the situation nodes at the time are calculated as shown in Table 5.

concentrating (30 minutes) → sleepy (15 minutes) → concentrating (20 minutes) → present

In learning phase, the system constructs the timing-structure model of human speech from the learning data. t_k is the value of a situation node s_j when a utterance c_i is observed. n is the total number of each utterance. The conditional probability $P_{s_j|c_i}(t)$ for an utterance node and a situation node are calculated as follows.

$$f_k(t) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(t-t_k)^2}{2\sigma^2}\right) \tag{4}$$

$$P_{s_j|c_i}(t) = \frac{\sum_{k=1}^{n} f_k(t)}{n} \tag{5}$$

In selecting phase, the probability of each utterance node is calculated using Equation 6 and the system utters a word when the probability exceeds the certain threshold and attains the local maximum point.

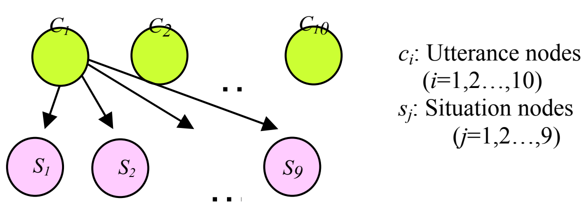


Fig. 4. Model of Bayesian network

Table 4. Situation nodes

The meaning of nodes	Number of nodes
Total time of a day of each state of consciousness	4
Duration of each state of consciousness	4
Elapsed time from the last utterance	1

Table 5. Example of situation nodes

Total time		Duration	
concentrating	: 50 minutes	concentrating	: 20 minutes
lack of drive	: 0 minutes	lack of drive	: 0 minutes
sleepy	: 15 minutes	sleepy	: 0 minutes
sleeping	: 0 minutes	sleeping	: 0 minutes

$$BEL(c_i) = P(S^i \mid c_i) = \prod_{j=!}^{m} P_{s_j \mid c_i}(s_j{}^i) \ . \qquad (6)$$

Each threshold is set to 1/2 of the minimum BEL in study data.

3 Evaluation Experiments

We conducted evaluation experiments for each part and the whole system.

3.1 Experiment of Posture Estimation

First, we conducted the experiment of posture estimation. We supposed that we can classify postures into four postures and we confirmed whether the system can estimate the posture.

Posture images are taken for ten minutes with 2 fps (1200 frames) and judged the posture per one frame. We prepared four types of postures “normal (upright) ”, ”leaning”, ”downward” and ”face down ”. To get answers, posture images are preserved and the user judges it. Posture estimation from the system is compared with the correct answer and recognition accuracy is calculated. This experiment is conducted at a desk in the laboratory and the USB camera is set on the personal computer.

The threshold value of Equation 1 is set to 20 from preliminary experiments. If we set a smaller threshold, image is influenced by the flicker of the fluorescent lamp of the room. If we set a larger threshold, colors in background cannot be recognized.

Results of the accuracy of posture estimation were as shown in Table 6. Results showed high recognition rates of each posture but the rates of “downward“ and “face down” are lower than that of the other postures. This is because downward or face down posture leaning a little is occasionally estimated as leaning posture.

3.2 Experiment of Consciousness Estimation

We conducted the experiment of consciousness estimation.

A user’s activity data is taken for 7 days and a system estimates the user’s consciousness per 5 minutes. Four levels of consciousness, ”concentrating”, ”sleepy”, ”lack the drive” and ”sleeping” are prepared. The estimation is compared with the correct answer that the user judges.

As a result, the accuracy of consciousness estimation was calculated as shown in Table 7. Results showed high recognition rates of each consciousness. However, when feeling a little tired even if concentrating, the state is sometimes estimated as lack of drive. Moreover, it was difficult to distinguish between “sleepy” and “lack of drive”. It was turned out that there were a few neutral states like this. To cope with these situations, it is one way that we make the system recognize that the judgment is inaccurate and prepare utterances corresponding to the situation.

Table 6. Accuracy of posture estimation

	Normal	Leaning	Downward	Face down	Total
Recognized frames (frames)	277	380	193	267	1117
Total number (frames)	287	382	229	302	1200
Accuracy (%)	96.5	99.5	84.3	88.4	93.1

Table 7. Accuracy of consciousness estimation

	Concentrating	Lack of drive	Sleepy	Sleeping	Total
Recognized number	146	9	10	10	175
Total number	148	9	11	10	178
Accuracy (%)	98.6	100	90.9	100	98.3

Table 8. Threshold of each utterance

	Utterance	Threshold($\times 10^{-13}$)
1	Are you ok?	0.098548
2	Wake up!	0.063544
3	You seem tired.	0.105729
4	Snap out of it.	0.038767
5	Do you want to sleep?	0.052691
6	Are you tired?	0.122829
7	Why don't you take a rest for a while.	0.292587
8	You started studying hard.	0.222411
9	You study hard!	0.002201
10	Don't work too hard.	1.279856

Table 9. Example of utterance corresponding to consciousness estimation (self-judge)

Consciousness state	Utterance from system
Sleeping	Are you ok?
Concentrating	You study hard!
Sleepy	Wake up!

3.3 Experiment of Utterance

We conducted the experiment of utterance part.

First, the teacher data are made by providing a teacher utterance for active data or unreal data and Bayesian network learns the utterance patterns. 67 data were used for its learning. Secondly, we calculated probabilities and decided thresholds of each utterance. Finally, we let users study at a desk while this system is running and observed utterances from the system. Software “Easy Speech” [9] is used for generating utterances.

We decided the threshold of each utterance as shown in Table 8. Each threshold is set to 1/2 of the minimum BEL in study data. Using these thresholds, we let the system running, and confirmed utterances as shown in Table 9. When concentrating, the transition of probability of “You study hard ! ” is as shown in Figure 5.

The threshold varied widely per utterance. This is because the probability of data far from other data becomes very small when the data increases and leans. For example, concerning utterance 9 “You study hard ! ” which has a lot of learning data, figure 6 shows the conditional probability of “Utterance 9” and “Duration of concentrating”. The probability after 100 minutes becomes extremely small where the number of teacher data is few. Therefore, it is difficult to use only threshold to judge an utterance timing and we put local minimum condition to the judgment method. Although the threshold of utterance 9 is 0.002201 and this is very small, it is expected that utterance is generated in both the first half part (35, 70, 100 minutes) when the probability is large and the latter part (140, 180, 220 minutes) when the probability is small.

Secondly, the utterances from the system were compared with consciousness state (self- judge) as shown in Table 9. The results show that the utterances are suitable to the situations like worry in case of sleeping and praise in case of concentrating. We also analyzed Bayesian probability. Figure 5 shows a time evolution of Bayesian probability. The probability of “You study hard!” increased gradually as concentrating time passed and the system uttered at its top. The reason of decrease after the last utterance is that passed time after utterance is one of situation nodes. The conditional probability between utterance 9 and passed time after the last utterance is as shown in Figure 7.

3.4 Effectiveness of Whole System

We confirmed the effectiveness of this system in daily life.

While this system is running, we let user lead a daily life in laboratory and observed utterances of the system. When the total of pixel values of user’s silhouette is less than 640*420*0.1=26880, it is judged that the user is not there and the system stops temporally.

Two people which are 1 female and 1 male participated in this experiment respectively for three days. The system was made by female’s teacher data. Table 10 shows the example of utterances. As a result, although there were some cases that no utterance was generated when long time passed, utterances were observed in any case within one hour. Moreover, we conducted a survey to the user and confirmed that the utterances were generated in appropriate circumstances.

The lack of the teacher data is thought as a reason why any utterance is not generated when long time passes. The variation of the transition increases as time passes, so the distance from teacher data becomes larger. To solve this problem, we should increase teacher data or make the variable of Bayesian network larger as time passes.

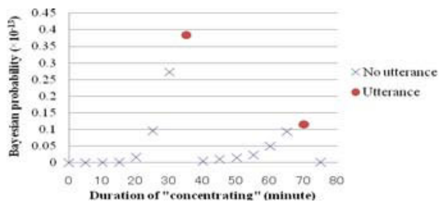


Fig. 5. Time-sequence data of Bayesian probability of “You study hard!” while concentrating

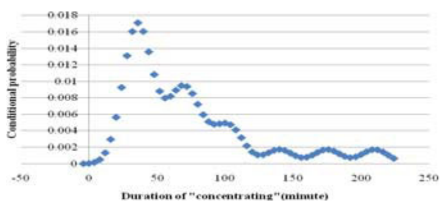


Fig. 6. Conditional probability of “Utterance 9” and “Duration of concentrating”

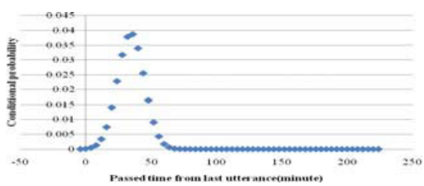


Fig. 7. Conditional probability of “Utterance9” and “Passed time from last utterance”

Table 10. Example of experimental result

Passed time (minute)	Consciousness state (system)	utterance
5	Sleepy	
10	Sleepy	
15	Sleepy	Wake up !
20	Lack of drive	
25	Lack of drive	
30	Concentrating	
35	Concentrating	You started studying hard!
40	Concentrating	
45	Concentrating	
50	Concentrating	
55	Concentrating	
60	Concentrating	
65	Sleepy	Wake up !

4 Conclusion

We developed a study support system as one of the spontaneous support system. This system got the awareness from user's appearance and talked to the user actively. We will try to focus on other motions and widen the scope of utterances in the future.

References

1. Okada, M.: Come Close to Me!: To be more Sociable Robots for Us. Journal of Human Interface Society 10(2), 87–92 (2008)
2. Oka, N., Kushi, K., Sonoyama, T., Shibuya, Y.: An Education in Interaction Design with Robot Materials. Journal of Human Interface Society 10(2), 105–108 (2008)
3. Ono, T., Ogawa, K.: Ubiquitous Mind and Interaction Design between Human and Robot. Journal of Human Interface Society 10(2), 115–118 (2008)
4. Kozima, H.: What Children Experience in Interactive Play with Robots. Journal of Human Interface Society 10(2), 99–104 (2008)
5. Ueda, H.: Human-Symbiotic Robots in the Home Ubiquitous Environment. Journal of Human Interface Society 10(2), 93–98 (2008)
6. Sato, T.: Networked Intelligent Machine to Support Daily Life —Trials of Robotic Room. Toshiba Review 56(9), 28–32 (2001)
7. Adachi, K.: Monitoring Car Drivers' Condition Using Image Processing. The Transactions of the Institute of Electrical Engineers of Japan. A Publication of Sensors and Micromachines Society 126(2), 31–37 (2006)
8. Hashimoto, K.: "Safety Human Engineering", Japan Industrial Safety & Health Association (1984) (in Japanese)
9. Easy speech, `http://www.vector.co.jp/soft/dl/winnt/art/se367527.html`

Author Index